



Università di Pisa

Dipartimento di Civiltà e Forme del Sapere

Corso di laurea Magistrale in Filosofia e Forme del Sapere

Tesi di laurea

Intelligenza artificiale: I primi 50 anni

Candidato

Marco Messina

Relatore

Prof. Claudio Pogliano

A.A. 2014/2015

Sommario

Introduzione	3
1. Dalla cibernetica all'AI.....	9
1.1 La psicologia cibernetica.....	9
1.2 Le macchine di Turing.....	18
1.3 Calcolatori digitali e primi linguaggi di programmazione.....	23
1.4 Calcolatori intelligenti.....	28
1.5 Simulare l'intelligenza nei calcolatori.....	34
1.6 Calcolatori e flessibilità.....	45
1.7 Decisioni complesse.....	49
2. Gli <i>Automata Studies</i>	57
2.1 Claude Shannon.....	58
2.2 John McCarthy.....	63
2.3 Il reclutamento per gli <i>Automata Studies</i>	67
2.4 Una falsa partenza	74
2.5 Il finanziamento per il seminario.....	88
3. La Conferenza di Dartmouth.....	98
3.1 Marvin Minsky	98
3.2 La Proposta di Dartmouth.....	104
3.3 Herbert Simon e Allen Newell.....	109
3.4 Il <i>Logic Theorist</i>	114
3.5 Prima, durante e dopo Dartmouth.....	122
3.6 Un congresso controverso	130
4. L'AI in Italia negli anni Cinquanta-Sessanta.....	143
4.1 <i>Adamo II</i>	143
4.2 1962	160
4.3 1967.....	174
5. <i>AI@50</i>	197
3.1 McCarthy e Minsky sullo stato dell'AI.....	201
3.2 Uno sguardo sul presente.....	207
3.3 Previsioni per il 2056	233

Introduzione

L'intelligenza artificiale (da qui in poi solo AI: Artificial Intelligence) ha come data di nascita ufficiale il 1956, anno in cui si tenne la famosa conferenza estiva presso il Dartmouth College di Hannover, nel New Hampshire. Fu qui che, per la prima volta, vennero raccolti i contributi sviluppati durante gli anni precedenti nel campo dell'AI simbolica. Questa data, sulla quale vi è accordo pressoché unanime nella comunità scientifica, potrebbe far pensare che la questione delle origini si possa considerare definitivamente risolta. In realtà non è così semplice. Quando si parla di storia dell'AI in genere non si è soliti partire dagli sviluppi successivi al 1956, ma si fa riferimento agli anni Quaranta del secolo scorso, alla Seconda guerra mondiale, alla cibernetica e all'avvento dei primi calcolatori digitali. E' senza dubbio in quel periodo che si assistette alla definitiva legittimazione scientifica dell'analogia uomo-macchina, condizione imprescindibile per l'avvento dell'AI: ciò sarà l'argomento principale del primo capitolo. Dalla fine degli anni Quaranta si cominciò a indicare con il termine cibernetica lo studio sistematico dei processi riguardanti la comunicazione e il controllo sia negli animali sia nelle macchine. Risale al 1943 quello che viene generalmente considerato il testo fondatore della futura disciplina, *Behavior, Purpose and Teleology*, a opera di Arturo Rosenblueth, Norbert Wiener e Julian Bigelow. L'idea alla base

del progetto cibernetico era quella di studiare i meccanismi dell'autoregolazione e del controllo presenti sia negli organismi viventi sia nelle macchine dotate di retroazione, in grado cioè di rispondere in modo adattativo alle sollecitazioni dell'ambiente esterno, modificando il proprio comportamento. Fu la Seconda guerra mondiale, con il suo singolare connubio tra scienze accademiche e sforzo militare, a fornire lo scenario ideale per l'affermarsi di queste idee. Nonostante l'analisi di Rosenblueth, Wiener e Bigelow partisse dai semplici servomeccanismi, essa finì con l'includere anche altri tipi di approcci, tra cui, in un primo tempo, l'AI simbolica. Ma se la prima cibernetica rimase sempre profondamente ancorata a una concezione tendenzialmente meccanicistica dell'essere umano, l'AI, di contro, poneva l'accento sulla manipolazione simbolica della conoscenza: da una parte il cervello come un hardware, dall'altra la mente come un software. Bisognerà aspettare la seconda metà degli anni Sessanta affinché questa incompatibilità di fondo venisse riconosciuta dai più.

Per potersi affermare, l'AI necessitava di un sistema artificiale nel quale riprodurre, emulandoli, i fenomeni dell'intelligenza, e il calcolatore digitale divenne il miglior candidato per questo compito. Alla base del calcolatore digitale vi è il concetto di macchina di Turing, un sistema in grado di assumere un numero finito di stati diversi e di eseguire un numero limitato di azioni, al fine di poter esprimere qualsiasi tipo di procedura definita. Nel

1950, Claude E. Shannon scrisse uno dei primi articoli sulla programmazione computerizzata, nel quale si faceva (seppur implicitamente) riferimento all'elaborazione simbolica. Sforzi simili vennero compiuti in quegli anni anche da Maurice Wilkes e Anthony G. Oettinger.

Nella tesi verranno prese in esame due conferenze, tenutesi rispettivamente nel 1955 e nel 1956 (il *Symposium on the Design of Machines to Simulate the Behavior of the Human Brain* e il *Symposium on the Impact of Computers on Science and Society*), durante le quali si affrontarono già alcune delle questioni che avrebbero caratterizzato il futuro (e le controversie) dell'AI, tra cui la rappresentazione della conoscenza, il ruolo della neurologia e la simulazione delle capacità cognitive tramite i calcolatori (sebbene applicata alle reti neurali).

Il secondo capitolo prende in esame la collaborazione tra Claude Shannon e John McCarthy, la quale avrebbe portato alla pubblicazione degli *Automata Studies*. La corrispondenza tra i due mostra come Shannon, fondatore della teoria dell'informazione, non fosse così conservatore come solitamente lo si dipinge in questa circostanza: egli condivideva con McCarthy lo scopo di creare macchine intelligenti, ma allo stesso tempo possedeva una visione d'insieme più ampia sull'argomento, la quale non includeva solo l'elaborazione simbolica, ma anche altri approcci tipicamente cibernetici, come le reti neurali e i sistemi a retroazione. Di contro, lo scopo di

McCarthy non era studiare e costruire macchine che replicassero, strutturalmente e funzionalmente, il cervello umano e i comportamenti intelligenti che scaturiscono dalle sue funzioni, quanto piuttosto progettare macchine in grado di esibire un comportamento che in certi contesti avrebbe potuto definirsi intelligente, prendendo i processi cognitivi come esempi di procedure efficaci: da qui il desiderio di prendere le distanze dalla cibernetica, vista come un approccio del tutto inadeguato allo scopo. L'insoddisfazione di McCarthy verso gli *Automata Studies* lo spinse a chiedere un finanziamento alla Rockefeller Foundation, al fine di organizzare una conferenza avente come oggetto l'AI. Questa sarebbe passata alla storia come la Conferenza di Dartmouth, e su di essa ci si concentra nel terzo capitolo.

Nell'estate del 1956 un gruppo di studiosi si riunì quindi al Dartmouth College con lo scopo di esaminare, come apparve nella proposta per il seminario redatta l'anno precedente, la congettura che ogni aspetto dell'intelligenza potesse essere, in linea di principio, descritto in modo tanto preciso da far sì che una macchina lo simuli. Gli organizzatori erano McCarthy, Marvin Minsky, ricercatore di matematica e neurologia ad Harvard, Nathaniel Rochester, direttore della ricerca sull'informazione in un centro ricerche dell'IBM, e Claude Shannon. Il seminario avrebbe dovuto avere le caratteristiche del *brainstorming*, ossia di un dibattito aperto e poco strutturato, ma i problemi organizzativi e la poca

collaborazione tra i partecipanti ne resero impossibile la perfetta riuscita. Fu durante il suo svolgersi che venne presentato il *Logic Theorist* di Allen Newell, Bernard Shaw e Herbert A. Simon, il primo programma euristico della storia, in grado di dimostrare teoremi della logica del primo ordine. Negli anni immediatamente successivi alla conferenza vennero raggiunti i primi significativi successi nel campo dell'AI, come il *GPS* di Newell e Simon e l'*Advice Taker* di McCarthy, un programma mai effettivamente realizzato ma di grande importanza concettuale. Ma furono i finanziamenti dell'ARPA (Advanced Research Projects Agency) a giocare un ruolo determinanti per il futuro dell'AI. Contemporaneamente, si assistette al declino dei modelli cibernetici, in particolare delle reti neurali, anche a causa della critica decisiva di Minsky al *Perceptron* di Frank Rosenblatt, un sistema neurale incapace di riconoscere stimoli visivi anche molto semplici. Percepita ormai come un approccio improduttivo, le risorse prima destinate alla cibernetica furono quasi completamente convogliate sull'AI. In un certo senso, quella tradizione risorgerà verso la metà degli anni Ottanta con il riemergere, in seno all'AI, del paradigma delle reti neurali.

Negli anni Cinquanta-Sessanta, l'Italia si dimostrò poco reattiva a questi mutamenti: pur accogliendo e facendo propri i contenuti della cibernetica, in quel periodo l'AI fu invece oggetto di studio di un numero estremamente limitato di ricercatori, i quali adottarono nei suoi confronti un punto di vista spesso molto critico, seppur non privo di obiezioni fondate e in anticipo sui

tempi. Nel quarto capitolo saranno presi in esame il lavoro del filosofo e linguista Silvio Ceccato e gli articoli pubblicati sull'*Almanacco letterario Bompiani* del 1962, dedicato alle applicazioni dei calcolatori elettronici alle scienze morali e alla letteratura. Saranno inoltre esaminate due importanti conferenze svoltesi nel 1967: il XXI Congresso della Società Filosofica Italiana, tenutosi a Pisa e avente come oggetto *L'uomo e la macchina*, e il congresso su *L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, organizzato dall'Accademia dei Lincei.

Nel 2006, per celebrare il cinquantenario della Conferenza di Dartmouth, il Dartmouth College organizzò una nuova conferenza, la *Dartmouth Artificial Intelligence Conference: The Next Fifty Years* (o *AI@50*), che sarà presa in esame nell'ultimo capitolo. Vi si mostrarono non soltanto i successi raggiunti dall'AI nell'intervallo tra i due eventi, ma anche tutti quegli aspetti controversi che da sempre hanno caratterizzato questa disciplina scientifica.

1. Dalla cibernetica all'AI

1.1 La psicologia cibernetica

Nel giugno del 1940, il presidente Roosevelt istituì una nuova struttura federale, il National Defense Research Committee (NDRC), al fine di reclutare tutti i centri di ricerca civile che si sarebbero potuti rivelare utili dal punto di vista militare. Ne era a capo Vannevar Bush, presidente del Carnegie Institution di Washington. Appena un anno dopo, il NDRC e il Committee on Medical Research si unirono dando origine all'Office of Scientific Research and Development (OSRD), sempre guidato da Bush. L'interdisciplinarietà era la modalità principale di questo nuovo tipo di ricerche, caratterizzate dalla stretta collaborazione tra scienziati, ingegneri, antropologi, sociologi, psicologi e militari.¹ Fu la Seconda guerra mondiale a offrire lo scenario ideale per la realizzazione di questo potenziale:

The war also opened up whole new areas of governmental activity as fields of inquiry for social scientists. Psychological warfare research, morale studies, and propaganda analysis became firmly established as specialized fields.²

In particolare, la psicologia fu tra le discipline a giovare maggiormente: tramite la guerra essa guadagnò visibilità e legittimazione, oltre che cospicui finanziamenti per la ricerca. Lo scopo della psicologia è sempre stato quello di rappresentare la “natura umana” come oggetto di conoscenza

¹ Pogliano 2004: 176.

² Crawford, Biderman 1969: 9.

scientifico e individuare le norme di comportamento. Ma esattamente come la quasi totalità delle scienze, anche le teorie psicologiche sono figlie del proprio tempo, e da esso influenzate. In questo senso, il secondo conflitto mondiale non fece eccezione.

Il lavoro di molti psicologi su problemi prettamente militari cambiò i loro interessi e la direzione delle loro ricerche. Gli strumenti di questa trasformazione furono politici, sociali e tecnologici oltre che teorici. La psicologia della Seconda guerra mondiale fu, per usare un'espressione di Paul Edwards, una "produzione militarizzata di conoscenza" (*militarized knowledge production*), nel senso che gli scopi militari finirono con il delinearne le direttive e i problemi particolari da affrontare, alla ricerca di soluzioni militarmente utili e pratiche per le esigenze contingenti.

Risultato di queste trasformazioni fu la nascita, tra gli anni Quaranta e Cinquanta, della "psicologia cibernetica", ovvero delle teorie proto-cognitive, il cui scopo era capire i processi della percezione, della memoria e del linguaggio nei termini di trasformazioni formalizzabili dell'informazione e di circuiti a retroazione. Essa nacque dal tentativo di teorizzare l'essere umano come parte costituente di un sistema bellico e continuò, anche dopo la guerra, a fornire modelli essenziali a questo scopo.³ E' in questo contesto che si inserisce il lavoro di Norbert Wiener. Agli inizi degli anni Trenta, egli ebbe modo di conoscere il neurobiologo messicano Arturo Rosenblueth, con il quale scoprì di condividere il forte interesse per

³ Edwards 2000: 340-343.

la metodologia scientifica e l'idea che la scienza dovesse essere “a collaborative effort”⁴, capace di trascendere la specializzazione delle singole discipline.

Verso la fine degli anni Trenta, Rosenblueth iniziò a tenere una serie di incontri mensili aventi come oggetto proprio il metodo scientifico e di cui Wiener divenne presto membro attivo. L'idea di fondo era quella che solo la dialettica tra ambiti di specializzazione diversi, tra scienziati “each possessing a thoroughly sound and trained acquaintance with the fields of his neighbors”, capace di unire matematica e psicologia, potesse riempire le “zone di confine” delle scienze e portare a un sapere unitario in questi campi inesplorati. Anche se la Seconda guerra mondiale ritardò l'attuazione pratica di questo sogno, essa fornì, come detto precedentemente, lo scenario ideale per la realizzazione e l'esecuzione di alcune delle sue caratteristiche basilari.

Durante il conflitto, il prestigio dell'aviazione tedesca e la posizione di difesa dell'Inghilterra indirizzarono l'attenzione di molti scienziati al potenziamento dell'artiglieria antiaerea. La velocità degli aerei aveva reso obsoleti tutti i metodi classici di puntamento, creando la necessità di servomeccanismi in grado di predire la posizione dei velivoli avversari, permettendo così il puntamento automatico delle armi e riducendo di conseguenza l'errore umano. Nel 1941 Wiener, al tempo professore di matematica al MIT, si unì a un gruppo di scienziati al Radiant Laboratory,

⁴ Wiener 1956: 171.

impegnati a studiare proprio questo problema sotto finanziamento del NDRC. Lavorando con Rosenblueth e l'ingegnere Julian Bigelow, Wiener iniziò a pensare al modo in cui sarebbe stato possibile predire la posizione futura di un aereo avendo come informazioni solo la posizione presente e la velocità, elaborando quindi una teoria di predizione statistica basata su informazioni incomplete. Era la teoria del controllo in retroazione, che sarebbe diventata la base per la progettazione dei servomeccanismi, e la cui originalità risiedeva nell'includere matematicamente il "fattore umano" (le caratteristiche operative sia di chi controlla l'artiglieria sia di chi guida l'aereo nemico) nello studio delle macchine che essi utilizzano.⁵

Nel 1942 la teoria crebbe abbastanza da convincere Rosenblueth a darne una presentazione estesa durante un breve seminario interdisciplinare a inviti, il Cerebral Inhibition Meeting, coordinato da Frank Fremont-Smith, direttore medico della Josiah Macy Jr. Foundation. Il seminario era centrato su due argomenti: l'ipnosi e la fisiologia dei riflessi condizionati (il fondamento della psicologia comportamentista).⁶ Questi, in quel periodo, erano argomenti di un certo interesse per la comunità scientifica: durante e dopo la Prima guerra mondiale, la suggestione ipnotica iniziò, infatti, a essere considerata un trattamento efficace per il cosiddetto trauma da bombardamento (*shell shock*), il corrispettivo maschile dell'isteria che durante il conflitto colpì un gran numero di soldati. Tra coloro che si

⁵ Wiener 1949: 7-11.

⁶ Fearing 1930: 1970; Hilgard, Marquis 1940: 1961.

interessarono del fenomeno vi fu Walter B. Cannon, fisiologo americano che ricondusse la causa del trauma a un abbassamento della pressione sanguigna. Nel 1930, appena arrivato ad Harvard, Rosenblueth divenne presto il braccio destro (per usare un'espressione di Wiener) di Cannon, il quale notò subito che il giovane messicano era particolarmente portato per l'indagine psicologica.⁷ Si trattava, insomma, di argomenti che Rosenblueth conosceva molto bene.

Tra gli altri partecipanti del seminario vi erano il neurologo Warren McCulloch, gli scienziati sociali Gregory Bateson, Margaret Mead e Lawrance K. Frank, oltre che lo psicanalista Lawrance Kubie. Con la promessa di un nuovo paradigma di ricerca interdisciplinare, la presentazione di Rosenblueth fu un autentico successo, tanto che da quel momento gli assunti della cibernetica iniziarono a diffondersi e affermarsi rapidamente all'interno della comunità scientifica.

La teoria venne presentata ufficialmente l'anno dopo con l'articolo *Behavior, Purpose and Teleology*. In sostanza l'idea di Rosenblueth-Wiener-Bigelow era quella di riconoscere in un'ottica comportamentista determinate attività degli organismi e analizzarle in termini di funzionamento di macchine analogiche. L'analisi tuttavia differiva in maniera rilevante dalle direttive del comportamentismo classico. Questo era solito concepire l'organismo come relativamente passivo all'ambiente, un *Black Block* che si limitava a reagire meccanicamente agli stimoli esterni.

⁷ Pogliano 2004: 172-174.

Ora, però, veniva introdotto il concetto di comportamento finalizzato a uno scopo. Spiegare un'azione in termini di risultato da raggiungere era stato da sempre un metodo criticato dagli scienziati in quanto avrebbe significato giustificare l'azione sulla base di qualcosa che non era ancora avvenuto: un sovvertimento causale ingiustificabile. Travalicando la secolare opposizione tra teleologia e meccanicismo, da quel momento si sarebbe potuto parlare esplicitamente di “meccanismi teleologici”.

I comportamenti teleologici orientati a uno scopo venivano descritti come movimenti controllati da retroazione negativa.⁸ Un oggetto tendente a uno scopo era costantemente aggiornato circa l'ambiente che lo circondava, ricevendo da esso informazioni sotto forma di input che a loro volta regolavano i suoi comportamenti futuri, il tutto prima che l'obiettivo venga raggiunto.

Nell'articolo venivano sottolineate le analogie tra servomeccanismi e i comportamenti degli esseri viventi guidati dagli organi di percezione.

Sostanzialmente, sono tre gli aspetti importanti da sottolineare: l'analisi comportamentista applicabile sia alle macchine sia agli esseri viventi; la rielaborazione di concetti psicologici e filosofici usando la terminologia tipica dell'ingegneria delle comunicazioni; l'analisi degli esseri umani quali componenti di un sistema bellico come fonte principale di analogie.

Nello scritto Rosenblueth, Wiener e Bigelow, infatti, contrapponevano comportamentismo e funzionalismo, scegliendo il primo sia perché non si

⁸ Heims 1994: 18-19.

soffermava sulle differenze strutturali tra macchine e organismi viventi, sia perché, presumibilmente, permetteva di applicare il formalismo matematico al lavoro di *radar tracking*.⁹

La psicologia cibernetica fu quindi inizialmente introdotta come un'estensione del comportamentismo. Ma come teoria generale di meccanismi auto-regolati, la cibernetica avrebbe trasceso le differenze tra macchine e organismi, e l'avrebbe fatto non rigettando concetti come comportamento, scopo, e volontà (come la psicologia comportamentista), ma ampliando la categoria di "macchina", tramite il concetto di retroazione, per includere queste nozioni. In realtà la nuova psicologia cibernetica si collocava a metà tra due tendenze, perché offriva nuovi strumenti di ricerca sia alle descrizioni fortemente matematizzate del comportamentismo sia a quelle concernenti i processi interni tipiche del formalismo. Non a caso nell'articolo si legge che:

In future years, as the knowledge of colloids and proteins increases, future engineers may attempt the design of robots not only with a behavior, but also with a structure similar to that of a mammal. The ultimate model of a cat is of course another cat, whether it be born of still another cat or synthesized in a laboratory.¹⁰

Sebbene l'analisi di Wiener partisse dai semplici servomeccanismi, essa finì con l'includere anche altri tipi di approcci e di meccanismi (tra cui, per un certo periodo, anche l'AI simbolica). Nel 1948, Wiener pubblicò il libro

⁹ Edwards 2000: 344.

¹⁰ Rosenblueth, Wiener, Bigelow 1943: 23.

Cybernetics; or Control and Communication in the Animal and the Machine, in cui cercava di legare insieme diversi indirizzi scientifici fino a quel momento più o meno indipendenti tra loro: il calcolo elettronico digitale, la teoria dell'informazione, i primi lavori sulle reti neurali, la teoria dei servomeccanismi e dei sistemi a retroazione, ma anche teorie legate alla psicologia, alla psichiatria, al processo decisionale e alle scienze sociali.

La teoria cibernetica diede all'analogia con le macchine una nuova fondazione teorica, creando nuove terminologie tecniche e sistemi di quantificazione (la teoria dell'informazione) per descrivere comportamenti auto-diretti sia nelle macchine sia nel pensiero umano. Invece di limitarsi a descrivere matematicamente i comportamenti osservati, la cibernetica analizzava modelli matematici, processi d'informazione, macchine "fisiche" ed esperimenti su umani e animali all'interno di un'unica cornice concettuale di ricerca. Wiener ipotizzava che tutti i comportamenti intelligenti fossero il risultato di meccanismi a retroazione: la stessa l'intelligenza non era altro che il risultato del ricevere ed elaborare informazioni. Questo spostamento di paradigma dall'energia all'elaborazione dell'informazione si sarebbe rivelato fondamentale per tutte le successive ricerche sull'AI.

Per l'aver definitivamente legittimato il parallelismo tra uomo e macchina (o, per meglio dire, dei processi cui questi sono soggetti), molti storici sono soliti vedere nella cibernetica il punto di origine della prima AI. La

cibernetica del dopo guerra restava infatti profondamente ancorata a una concezione prettamente bio-meccanicistica. Per i primi cibernetici la macchina formale, il modello logico era, prima e innanzitutto, una macchina, un dispositivo a proprietà fisse regolato da un sistema di input e output. Il principale metodo applicativo della cibernetica rimaneva quello della costruzione di strutture complesse partendo da piccole unità fisiche come i neuroni e gli archi riflessi. La concezione principale restava quella della mente “incarnata”.

Generalizzando, si potrebbe dire che, mentre il principale scopo della cibernetica era quello di riprodurre il cervello umano in un hardware, l’AI puntava invece a simulare la mente in un software. Ma il passaggio dal modello biologico a quello simbolico, cui la Conferenza di Dartmouth viene generalmente considerata come ideale punto di rottura, fu in realtà molto più graduale. Con la sola possibile eccezione di John McCarthy, tutti i fondatori dell’AI erano, in una maniera o nell’altra, influenzati dal modello biologico della cibernetica. Anche dopo il 1956 la separazione tra *brain modelling* e approccio simbolico appariva tutt’altro che definitiva.

I motivi di questo spostamento di paradigma erano imputabili più a cause di origine pratica e politica che a un’elaborazione puramente teorica e “astratta”: fu uno di questi casi in cui la sperimentazione precedeva la premessa teorica. Secondo Paul Edward, la generazione successiva ai primi cibernetici, coloro che erano stati studenti di McCulloch, Pitts, Shannon e

Wiener, portò l'analogia uomo-macchina su un altro livello, concentrandosi sull'incorporeo, sull'astratto, sul pensiero piuttosto che sul cervello. Alcuni di loro erano troppo giovani per aver contribuito in prima persona al conflitto mondiale, e per aver vissuto quel clima unico che combinava le scienze accademiche con l'ingegneria militare; per questa generazione i calcolatori erano solo dei modelli matematici automatizzati dal forte fascino intellettuale.¹¹

La chiave per comprendere la teoria dell'intelligenza, per loro, risiedeva nella nozione di processo simbolico. Questa idea si fece largo come conseguenza dei primi tentativi di programmazione dei calcolatori digitali: uno sviluppo già previsto da Turing.

1.2 Le macchine di Turing

Volendo riassumere brevemente l'apporto di Alan Turing alla teoria delle macchine intelligenti, non si può prescindere da tre dei suoi scritti principali: *On Computable Numbers with an Application to the Entscheidungsproblem*, *Intelligent Machinery* e *Computing Machinery and Intelligence*.

Nel primo, Turing attuava una vera e propria vivisezione dei processi cognitivi umani, analizzando un uomo intento nei processi di calcolo.

L'assunto di base era che, di fronte a un calcolo, la mente umana procedeva

¹¹ Edwards 1996: 240-241.

per passaggi successivi. Tutto stava nell'isolare questi passaggi, che avrebbero dovuto essere semplici, piccole particelle di pensiero, e trovare un linguaggio per insegnare alla macchina un modo per affrontarli; in altre parole, programmarla. In questo modo Turing finì con il definire il calcolo come un processo di manipolazione meccanica di simboli.

Turing immaginava la sua macchina come un nastro infinito suddiviso in caselle contenenti un simbolo o uno spazio vuoto, che scorreva all'interno di un dispositivo capace di leggere, scrivere o sostituire i simboli. Uno strumento siffatto poteva eseguire quattro azioni meccaniche: muovere il cursore a destra o a sinistra, scrivere un simbolo o cancellarlo. Secondo Turing, qualsiasi operazione complessa poteva essere affrontata ricorrendo a queste quattro, semplici azioni. Il passo successivo consisteva nel somministrare alla macchina la sua "tavola di comportamento", una stringa di simboli attraverso la quale essa leggeva i passi dell'algoritmo da seguire, ovvero tutti quegli stadi intermedi che necessariamente dovevano essere affrontati prima di arrivare alla conclusione del problema sottoposto.

Secondo Turing, un calcolatore di questo tipo, programmato per replicare gli stessi passaggi discreti di una persona intenta in un'elaborazione simile, avrebbe svolto un'operazione analoga, per non dire identica, alla sua controparte umana. Molti degli elementi chiave dei calcolatori digitali erano già presenti in questa macchina puramente ipotetica: logica binaria (i quadrati possono essere bianchi o segnati), programmazione (con il nome di

“configurazione”, che determinava il comportamento della macchina ed era il risultato della combinazione tra la condizione attuale della stessa e del simbolo esaminato), memoria fisica (il nastro) e le operazioni di base di lettura e scrittura.¹²

Erano state gettate le basi teoriche della computazione digitalizzata: la maggior parte dei calcolatori digitali che vennero costruiti negli anni successivi, per quanto molto diversi, dal punto di vista della progettazione, dal primitivo lettore di nastri di Turing, erano di fatto delle macchine di Turing, i suoi eredi spirituali.¹³

Nonostante tutto, *On Computable Numbers* ebbe ben poca risonanza negli anni successivi la sua pubblicazione. Molti che lessero l’articolo di Turing lo fecero solo quando il loro lavoro era ormai avviato.¹⁴ L’unica eccezione fu John von Neumann, che incontrò Turing a Cambridge, e che divenne l’unico tra i primi pionieri nello studio dei calcolatori ad analizzare le analogie cervello-macchina in maniera significativa. Invece, la prima vera influenza di Turing sul pensiero americano si ebbe con McCulloch e Pitts, con il loro tentativo di applicare “la definizione di computabilità di Turing” allo studio delle funzioni neurali.¹⁵

Per quanto significativa, quella riguardante le analogie tra macchine ed esseri umani era solo una piccola e collaterale parte di *On Computable*

¹² Ivi.: 242.

¹³ Shannon 1953: 704.

¹⁴ Augarten 1984: 145.

¹⁵ Copeland 2004: 407.

Numbers, ma nei vent'anni successivi queste idee finirono con l'occupare una parte preponderante nelle ricerche di Turing. Durante la Seconda guerra mondiale egli ebbe la possibilità di costruire la sua ipotetica macchina a Bletchley Park. Nel 1948 scrisse un rapporto per il British National Physical Laboratory intitolato *Intelligent Machinery*. Questo scritto utilizzava largamente la stessa concezione comportamentistica del cervello usata da McCulloch, Pitts e dai cibernetici in generale. Il cervello veniva descritto come una macchina auto-organizzata che necessitava dell'insegnamento per migliorare il suo comportamento fino a renderlo intelligente. L'idea era quella che il cervello, ossia la macchina, dovesse "riuscire a farcela in un modo o nell'altro", e cioè che non dovesse acquisire la capacità di pensare grazie all'intervento di un essere superiore che la programmi rigidamente. Doveva esserci un modo in cui la macchina potesse imparare da sola, proprio come il cervello umano. Secondo questa concezione, l'intelligenza non veniva installata nel corpo fin dalla nascita come un impianto elettrico: sebbene i riflessi non condizionati (come la respirazione o la vista) siano dovuti a certe strutture del cervello, altre attività non lo erano. Prendendo l'esempio di due persone che parlano una lingua diversa, Turing ipotizzava che vi fosse una sorta di "zona franca" del cervello le cui funzioni fossero altamente indeterminate e plasmate dall'esterno. La corteccia del neonato veniva paragonata a una macchina disorganizzata, ma potenzialmente organizzabile per mezzo di adeguate

interferenze educative.

I processi educativi che aveva in mente erano quelli di una pedagogia fondata sul sistema del bastone e della carota. Una macchina programmata per eseguire operazioni alla lettera avrebbe sicuramente compiuto il proprio dovere, ma alla stregua di una persona senza buon senso che obbedisce ciecamente agli ordini, mentre per definirsi davvero intelligente avrebbe dovuto possedere il dono dell'iniziativa.¹⁶ L'attenzione di Turing sul concetto di apprendimento colloca questo scritto a metà tra le istanze cibernetiche, nel suo dare alla macchina una forma quasi biologica, e il metodo della programmazione simbolica, più vicino a ciò che l'AI sarebbe diventata. Si tratta di una specie di passaggio di transizione tra il *brain modelling* della cibernetica e i processi simbolici dell'AI.

Ma per il futuro della psicologia e dell'AI, il lavoro più importante di Turing è sicuramente *Computing Machinery and Intelligence*, pubblicato nel 1950. Qui per la prima volta compariva il celebre Test di Turing, o gioco dell'imitazione, introdotto per presentare una formulazione alternativa alla classica domanda “possono le macchine pensare?”. Il gioco richiedeva tre partecipanti: un uomo (A), una donna (B) e l'interrogatore (C). Separato da una porta dagli altri due partecipanti, l'interrogatore doveva capire chi fosse l'uomo e chi la donna, con il primo intento a depistarlo e la seconda ad aiutarlo. Ma se si sostituisse A con un macchina, allora “the interrogator decide wrongly as often when the game is played

¹⁶ Hodges 2006: 491-494.

like this as he does when the game is played between a man and a woman?”.¹⁷ Il problema così posto aveva il vantaggio di tirare una separazione piuttosto netta tra le capacità fisiche e quelle intellettuali di un essere umano, concentrandosi quasi del tutto su quest’ultime quale discriminine per l’intelligenza, e isolando i parallelismi biologici tipici della cibernetica. Alla macchina non si richiedeva solo di elaborare numeri e nozioni logiche, ma anche il linguaggio scritto, in modo da relazionarsi con essa a un livello di quasi parità. Il calcolatore diventava ben più che un semplice elaboratore: era il processore simbolico per eccellenza, una macchina universale d’informazione.

Nessuna macchina ai tempi poteva neanche lontanamente eseguire quanto richiesto dal test. Né Turing dava alcun indizio su come avrebbe potuto. Nonostante ciò, l’idea di poter comunicare con un calcolatore tramite il linguaggio ordinario influenzò le ricerche per tutto il decennio successivo.

1.3 Calcolatori digitali e primi linguaggi di programmazione

Per aver posto l’accento sulle analogie di carattere organizzativo e funzionale tra cervello e calcolatori, si può senza dubbio definire Turing come il progenitore intellettuale dell’AI. Tuttavia, sembra che nell’immediato la sua influenza sui ricercatori americani e inglesi sia stata quasi nulla. Sebbene con il senno di poi possa sembrare ovvio che i neonati calcolatori digitali erano gli strumenti più adatti per realizzare il suo

¹⁷ Turing 1950: 433.

progetto, tra il circolo dei cibernetici vi era molto scetticismo circa la possibilità di riprodurre tramite questi dispositivi sia l'abilità di apprendere sia la complessità e la spontaneità del comportamento degli organismi viventi. Per loro, i comportamenti pre-programmati dei calcolatori peccavano di un'eccessiva rigidità.¹⁸ Traslando la domanda a là Turing, ci si potrebbe allora chiedere perché, a partire dagli anni Cinquanta, molti scienziati iniziarono comunque a domandarsi non tanto se fosse possibile riprodurre l'intelligenza in una macchina, quanto se fosse possibile farlo in un calcolatore, cioè in un dispositivo meccanico piuttosto specifico. L'ampia diffusione di cui i calcolatori furono protagonisti iniziò solo a partire dagli anni Cinquanta. Fino al 1950 vi erano pochi calcolatori esistenti, e solo poche centinaia di programmatori. Molti di essi non si identificavano come tali, dal momento che si occupavano anche di progettazione e costruzione. Ma già nel 1955 gli esemplari salirono a qualche migliaio, richiedendo l'impiego di circa 10.000 programmatori. Appena cinque anni dopo, nel bel mezzo del loro boom commerciale, programmare divenne un mestiere a tutti gli effetti, con circa 60.000 specializzati al lavoro su 5.000 calcolatori. Questo maggiore impegno nella costruzione e nella diffusione dei calcolatori negli Stati Uniti e in Europa fu fortemente sponsorizzato dal governo e dall'industria, tramite politiche che vedevano emergere e prevalere, forse per la prima volta, esigenze applicative di tipo amministrativo e commerciale, pur senza mai trascurare

¹⁸ Cordeschi 2002: 170.

le università, gli istituti di ricerca e le implicazioni militari: era il passaggio da una fase che potremmo definire “accademico-militare”, che aveva caratterizzato il Secondo conflitto mondiale, a una di tipo “commerciale”.¹⁹ A partire dalla metà del 1950, ad esempio, il SDC (System Development Corporation) iniziò ad assumere e addestrare migliaia di nuovi programmatori. Insegnati di musica e donne senza nessuna specializzazione erano tra le reclute più di successo.²⁰ Da un siffatto gruppo non ci si poteva certo aspettare che imparassero i codici macchina o che realizzassero eleganti algoritmi matematici: per rendere questa nuova forza lavoro efficace erano necessari linguaggi simbolici di programmazione facilmente apprendibili da non specialisti. Il bagaglio di conoscenze matematiche richiesto per queste operazioni era sempre più basso: presto ci si accorse che i calcolatori digitali avevano il doppio vantaggio di non richiedere una preparazione tecnica estremamente specialistica (almeno, non nei termini in cui la si intendeva prima) per essere utilizzati, e di permettere nuove possibilità d’impiego. L’incremento esponenziale della domanda di software, di programmatori non specializzati, di interfacce più *user-friendly* per motivi commerciali, e i maggiori investimenti economici sui calcolatori contribuirono alla nascita dei primi linguaggi simbolici di programmazione di alto livello.

Il primo di essi a essere disponibile per il mercato fu il FORTRAN,

¹⁹ Morelli 2001: 286-287.

²⁰ Tropp 1983: 386.

sviluppato nel 1954 dai ricercatori dell'IBM e rilasciato nel 1957. L'idea alla base del FORTRAN venne da John Backus, nel momento in cui si rese conto che le elevate potenzialità dell'IBM 704 erano drammaticamente ridimensionate dalle difficoltà della sua programmazione: ciò si traduceva in programmi scarsamente efficienti e, quindi, in una minore operatività della macchina. Di fronte a questa situazione, l'intuizione di Backus fu che soltanto un compilatore avrebbe potuto migliorare lo stato delle cose.²¹ E' necessario precisare però che, a quel tempo, gli algoritmi scritti in un linguaggio di programmazione di alto livello non erano visti come veri programmi, quanto piuttosto come indicazioni per il compilatore affinché componesse il programma: l'indipendenza del livello simbolico non aveva ancora raggiunto quello status assiomatico che gli sarebbe stato riconosciuto in seguito.

Resta comunque il fatto che, con l'introduzione del FORTRAN, chiunque avesse problemi di carattere scientifico poteva, senza difficoltà, tradurli in termini di espressioni (*statement*) con le quali programmare il funzionamento operativo del calcolatore; ciò che rendeva particolarmente prezioso l'impiego di questo linguaggio era la vicinanza di queste espressioni, come struttura logica e aspetto, alle stesse espressioni che descrivevano il problema con il quale ci si stava confrontando.²²

Dopo il FORTRAN fu la volta del COBOL, diretta conseguenza della

²¹ Un compilatore è un programma che provvede alla traduzione automatica delle istruzioni dai linguaggi simbolici di alto livello al linguaggio macchina.

²² Morelli 2001: 354-355.

Conference on Data System Languages (CODASYL). Caratteristica essenziale del COBOL, tuttora presente in molte applicazioni software commerciali di tipo bancario, è che esso può essere classificato come un linguaggio di programmazione di tipo discorsivo: fa infatti uso di verbi (in inglese) come operatori, e di nomi estesi come operandi. Il COBOL rappresentò una vera e propria rivoluzione nel modo di fornire alle macchine le istruzioni per il loro funzionamento, in quanto si focalizzava sulla logica dei problemi di natura amministrativa, contabile e finanziari, nonché sulle regole che presiedevano alla loro soluzione, rendendo più facile l'utilizzo dei calcolatori da parte di chi non aveva conoscenze specifiche riguardo il loro funzionamento e la loro architettura.²³

Ma mentre i programmatori sviluppavano linguaggi di programmazione *user-oriented* per scopi puramente pratici, il processo di creazione di questi portò alla consapevolezza del potenziale dei calcolatori come manipolatori non solo di numeri, ma anche di simboli di ogni tipo. Mentre Turing aveva avanzato la possibilità teorica di comunicare con i calcolatori tramite un linguaggio naturale, di fatto fu il lavoro pratico sugli stessi, nella creazione di software, che portò ai primi linguaggi di programmazione simbolici. Per molti dei fondatori dell'AI, l'idea dei calcolatori come incarnazione di un sistema universale di rappresentazione, un "linguaggio del pensiero"²⁴,

²³ Ivi.: 357-359.

²⁴ Fodor 1975.

emerse non in maniera puramente astratta e teorica, ma in quanto conseguenza di un'esperienza diretta con le macchine, sia come utenti sia come sviluppatori di linguaggi di programmazione di alto livello.²⁵

Quasi tutti i partecipanti della Conferenza di Dartmouth erano figli di questo particolare momento storico, ed è possibile inquadrare i temi portanti del loro manifesto programmatico all'interno del dibattito scientifico e culturale degli anni immediatamente precedenti la conferenza stessa.

1.4 Calcolatori intelligenti

Già nel suo articolo *Programming a Computer for Playing Chess* del 1950, Claude Shannon, uno dei quattro promotori della Conferenza di Dartmouth, riassumeva molto bene quali erano le caratteristiche concettuali in base alle quali molti avevano iniziato a considerare i calcolatori digitali come macchine potenzialmente intelligenti, nonostante le evidenti differenze strutturali rispetto agli esseri umani, e a rendere la questione di un certo interesse per la comunità scientifica. Del resto, Shannon aveva avuto modo di conoscere Turing, il quale gli fece leggere *On Computable Numbers*, ed entrambi avevano potuto constatare la loro convergenza di vedute circa la possibilità di riprodurre l'intelligenza umana in una macchina.²⁶

Con i calcolatori *general purpose* si disponeva per la prima volta di una macchina capace di manipolare strutture di simboli che il programmatore faceva corrispondere in modo naturale alle entità più diverse: parole di una

²⁵ Edwards 1996: 250-251.

²⁶ Hodges 2006: 327-329.

lingua naturale, espressioni matematiche, posizioni del gioco degli scacchi, oggetti da riconoscere e classificare. Shannon sottolineava proprio questa capacità di elaborazione simbolica quale elemento di assoluta originalità rispetto ai calcolatori precedenti. Un'altra delle loro caratteristiche fondamentali era quella di ricorrere a dei principi generali per raggiungere qualcosa di simile a un giudizio, scegliendo tra diverse alternative in base ai risultati precedentemente ottenuti: era la procedura del *trial-and-error*, secondo Shannon preferibile a quello dell'elaborazione rigida e inalterabile. Le soluzioni a questi problemi non potevano essere classificate in termini di giusto e sbagliato, ma essere disposte in uno spettro ascendente di "qualità" che andava dalla migliore possibile fino alla peggiore. In problemi complessi come il gioco degli scacchi, in cui le possibili alternative sono astronomiche, era necessario disporre di un "buon filtro", non necessariamente ottimo, ma in grado di dar comunque luogo a soluzioni soddisfacenti.²⁷ Queste idee facevano da sfondo alla Proposta di Dartmouth, e saranno giudicate importanti per tutto il decennio successivo.²⁸

Le ultime due caratteristiche cui Shannon alludeva (*trial-and-error* e la progressione per risultati soddisfacenti) erano possibili programmando il calcolatore digitale in base ad una serie d'istruzioni per cui, se certe condizioni erano soddisfatte, allora esse sarebbe state eseguite. Il "programma" consisteva in una serie di ordini elementari che implicavano

²⁷ Shannon 1950: 637.

²⁸ Cordeschi 2007: 261.

delle decisioni: se non vi erano le premesse, allora un altro ordine sarebbe stato eseguito, secondo un sistema d'istruzioni di salto condizionato.²⁹

Un calcolatore digitale in grado di eseguire queste operazioni era già stato creato appena un anno prima: si trattava dell'EDSAC, il primo grande calcolatore con "programma memorizzato"³⁰, progettato dal gruppo di Maurice Wilkes al Cambridge Mathematical Laboratory. Per esso Anthony Oettinger scrisse due dei primi programmi "intelligenti", in grado di modificare il proprio rendimento in base ai risultati ottenuti, mostrando una certa, per quanto primitiva, capacità di apprendimento.

Uno di essi, lo *Shopping Program*, simulava il comportamento di un bambino intento a fare compere. In un mondo virtuale, il bambino girava casualmente tra i negozi in cerca dell'articolo cercato e, una volta trovato, immagazzinava l'informazione per non dover ripetere da capo la ricerca la volta seguente. Il programma era in grado di manifestare una certa "curiosità", registrando la posizione anche di articoli non richiesti, in modo da trovarli subito in caso di necessità. Anche questo programma, nella sua elementare capacità di apprendimento, era basato sulle istruzioni di salto condizionato.³¹

Gli studi di Oettinger erano stati influenzati da tre articoli. Di due di essi si

²⁹ Shannon 1950: 645.

³⁰ Nella sua memoria interna erano infatti depositati non solo i dati, ma anche le istruzioni per manipolarli, ovvero il programma, che poteva essere modificato non meno dei dati mediante operazioni algoritmiche.

³¹ Cordeschi 2007: 262.

è già detto, mentre il terzo era *Can Machines Think?* di Wilkes.³²

In esso, Wilkes faceva riferimento al gioco dell'imitazione di Turing affermando che, per poter davvero simulare il comportamento umano tramite un calcolatore, sarebbe stato necessario progettare un programma di apprendimento generalizzato, cioè in grado di apprendere ogni soggetto scelto dal programmatore. Un obiettivo invero abbastanza fuori portata per i programmi del tempo.³³

Oettinger affermava che i suoi programmi soddisfacevano, almeno in parte, i requisiti posti da Turing e Wilkes. Molto lontani dal manifestare quell'abilità di apprendimento generale indicata da Wilkes, questi programmi erano comunque in grado di svolgere i loro compiti in vista di obiettivi specifici e molto ben definiti. Sarebbero stati, insomma, in grado di superare il gioco dell'imitazione in forma più ristretta e "soft". Era più che normale che i protagonisti della prima AI si concentrassero su problemi relativi ad ambiti ben delimitati, per risolvere i quali bastavano poche regole esplicite per l'elaborazione simbolica e poca conoscenza specializzata, visti i limiti dei calcolatori del tempo. L'EDSAC di Oettinger, ad esempio, peccava di eccessiva lentezza nella lettura dei nastri: sarebbe stato sicuramente più comodo fare in modo che la macchina, nella sua interezza, si occupasse esclusivamente dell'esecuzione del programma, del

³² Si trattava di *Computing Machinery and Intelligence* di Turing e, appunto, di *Programming a Computer for Playing Chess* di Shannon.

³³ Wilkes 1953: 1231.

bambino computerizzato (*child machine*), con il mondo esterno codificato da un nastro immesso nella macchina stessa. Ma ciò non era possibile, e risultava quindi più conveniente dividere l'EDSAC in due parti, designando parte dalla memoria al “soggetto sperimentale” e quella che restava alla rappresentazione dell'ambiente in cui interagire.³⁴

Il merito più importante di Oettinger stava sicuramente nell'aver evidenziato che i “digital computers can be made to serve as models in the study of the functions and of the structures of animal nervous systems”, un tema che avrebbe potuto interessare non solo coloro i quali si occupavano delle loro applicazioni non-numeriche, ma anche psicologi e neuropsicologi.³⁵ Oettinger sottolineava infatti che:

It is important to distinguish carefully between machines intended to be models of the physical structure of animal nervous systems and machines designed to perform specific functions of animal nervous system.³⁶

Lo stesso gioco dell'imitazione implicava solo una corrispondenza funzionale, e non strutturale, tra calcolatore e sistema nervoso. Si sarebbe anche potuto costruire una “macchina speciale” (*special machine*), la cui struttura fisica e processi interni non avrebbero avuto nessuna relazione con l'EDSAC e, in definitiva, con il cervello umano, se non per le corrispondenze superficiali riguardanti il sistema di input e output. Ma

³⁴ Oettinger 1952: 1247.

³⁵ Ivi.: 1243-1244.

³⁶ Ivi.: 1261.

l'EDSAC aveva il vantaggio non da poco, quando provvisto di un programma “adeguato” (*suitable*), di trasformarsi da “an orthodox computing machine into an experimental learning device”, tramite l’elaborazione simbolica: un obiettivo forse “markedly different from those its designers had in mind”, ma invero meritevole di approfondimento. Si può concludere dicendo che Oettinger, descrivendo le caratteristiche dei suoi (piuttosto semplici) programmi, fu tra i primi a sollevare alcuni punti che nell’immediato futuro sarebbero stati molto dibattuti tra filosofi, psicologi e ricercatori dell’AI. Inoltre la sua personale definizione di macchina intelligente in connessione con un programma “capable of performing function which, in living organisms, are considered to be the result of intelligent behavior”³⁷ sembrava anticipare la definizione di AI data dagli autori del documento preparatorio della conferenza di Dartmouth:

For the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving.³⁸

³⁷Ivi.: 1251.

³⁸McCarthy et al. 1955: 7.

1.5 Simulare l'intelligenza nei calcolatori

L'articolo di Wilkes fu ristampato in *Computer Issue*, un numero speciale dei *Proceedings of the IRE* (Institute of Radio Engineers), rivista scientifica pubblicata con la collaborazione del PGEC (Professional Group on Electronic Computers dell'IRE), nell'ottobre del 1953. A esso seguiva quello di Shannon intitolato *Computers and Automata*, che si soffermava ancora una volta sulla capacità dei calcolatori di eseguire operazioni non-numeriche, associabili alle funzioni del cervello umano. Seguiva una rassegna di alcune delle ricerche nel campo del calcolo non numerico e dei problemi che esse presentavano. In particolare, venivano sottolineati i limiti delle reti neurali nella riproduzione delle funzioni cerebrali:

The similarities between the brain and computers have often been pointed out. The differences are perhaps more illuminating, for they may suggest the important features missing from our best current brain models.³⁹

Qui Shannon sembrava suggerire la necessità di spostarsi verso altri indirizzi di ricerca fino a quel momento poco battuti. L'organizzazione logica, tra le altre cose, sembrava essere il vero discrimine tra calcolatori e cervello, poiché i primi, sprovvisti di un programma adeguato, sarebbero stati come degli *idiot savant*, persone incolte ma capaci di fare mentalmente lunghi calcoli aritmetici:

³⁹ Shannon 1953: 704

For long chains of arithmetic operations a digital computer runs circles around the best humans. When we try to program computers for other activities their entire organization seems clumsy and inappropriate.⁴⁰

Siamo nel pieno clima di diffusione dei calcolatori illustrato precedentemente; in Inghilterra Ferranti stava completando il Mark 1, mentre l'IBM iniziava a produrre il modello 701, descritto proprio in *Computer Issue* negli articoli di Werner Buchholz, Clarence E. Frizzell e Harold D. Ross. Si trattava del primo di una serie di calcolatori “generalisti” (*general purpose*) a programma memorizzato che sarebbero stati usati sia a fini teorici per la ricerca sia per applicazioni governative e industriali. Come ricercatore dell'IBM, Nathaniel Rochester, altro organizzatore della Conferenza Dartmouth, era responsabile dell'organizzazione logica del 701, e per esso scrisse il primo programma simbolico assemblativo.⁴¹ Nel 1952, il primo programma per giocare a dama, a opera di Arthur Samuel, l'autore dell'articolo di apertura di *Computer Issue* (oltre che uno dei futuri partecipanti del congresso), era già operativo su questo calcolatore.⁴² Questi e altri programmi erano illustrati da Shannon nel suo articolo. Nel 1954, Samuel progettò per l'IBM 704 il primo programma ad “auto-apprendimento” per giocare a dama, in modo che la macchina potesse imparare dalle proprie esperienze.⁴³ Da lì a breve, Newell e Simon

⁴⁰ Ibid.

⁴¹ https://www-03.ibm.com/ibm/history/exhibits/701/701_team.html

⁴² Wiederhold, McCarthy, Feigenbaum 1990.

⁴³ https://www-03.ibm.com/ibm/history/history/year_1956.html

avrebbero iniziato a lavorare al loro *Logic Theorist*. Oliver Selfridge, altro partecipante alla conferenza, fu tra i primi a simulare tramite i calcolatori le attività percettive dell'essere umano, sviluppando quello che poi sarebbe diventato il Pandemonium.⁴⁴ Una serie di simulazione delle reti neurali tramite calcolatori, derivanti dal lavoro seminale di McCulloch e Pitts, erano in fase di sviluppo, in particolare a opera di Farley e Clark (che nel 1961 simularono il comportamento di 1024 neuroni a connessioni casuali) e di Rochester insieme ad alcuni collaboratori, a partire dalla teoria di Donald Hebb sull'apprendimento e la formazione dei concetti. In più, Minsky e McCarthy si stavano già concentrando su diverse questioni concernenti l'AI.

Questi esperimenti sono menzionati nel documento preparatorio della Conferenza di Dartmouth, il quale, a riguardo delle reti neurali, ammetteva quasi placidamente che “partial results have been obtained but the problem needs more theoretical work”⁴⁵, mentre Hebb veniva citato solo indirettamente.

Ma un altro importante evento si tenne proprio in quel periodo: il simposio *The Design of Machines to Simulate the Behavior of the Human Brain*, sponsorizzato dal PGEC e tenutosi tra il 21 e il 24 marzo 1955. Gli interventi furono di McCulloch, Oettinger, Rochester e Otto Schmitt, biologo e figura eclettica, mentre John Mauchly, Marvin Minsky, Walter

⁴⁴ Cowan 1999.

⁴⁵ McCarthy et al. 1955: 2.

Pitts e Morris Rubinoff figuravano tra gli invitati.

Furono molte le questioni sollevate durante il simposio: computazione analogica contro digitale; la creatività dei calcolatori; il riconoscimento delle figure tramite gli stessi (durante il suo discorso, Rochester descrisse i risultati degli ultimi esperimenti di Selfridge, che non ebbe modo di partecipare); e la memoria “distribuita” (in una breve discussione con McCulloch, Minsky esprimeva scetticismo per questo modello, negando che un buon esempio ne potesse essere l’omeostato di Ashby, il quale “has no memory at all of its previous states, distributed or otherwise”⁴⁶).

Ma ai fini del nostro discorso è sicuramente di maggior interesse l’intervento su *Constraints and Similarities*, in cui Oettinger, come già nel suo articolo del 1952, distingueva tra due approcci diversi nella simulazione delle funzioni cerebrali tramite i calcolatori, che “although related, are far from being identical”.⁴⁷ Vi erano due possibili vie da percorrere, entrambe convergenti verso una maggiore comprensione del comportamento del cervello: una che, per così dire, seguiva l’esempio della natura, concentrandosi sulle analogie strutturali; e un’altra che attraversava strade che essa non aveva mai percorso.

Oettinger affermava che in un primo momento fosse più che naturale soffermarsi sulle funzioni organiche cercando di imitare ciò che si

⁴⁶ McCulloch et al. 1956: 251.

⁴⁷ Ivi.: 241.

conosceva o si ipotizzava sulla loro struttura, ma che non ci si dovesse soffermare solo su quello:

It should not be overlooked that many of the most successful simulations of living functions and, indeed, superior performances of these functions, have often been archived by means not generally used by living organisms.⁴⁸

Usando un esempio che sarebbe diventato celebre presso gli studiosi dell'AI, Oettinger affermava che “while the flight of birds undoubtedly stimulated man’s urge to fly, human flight was achieved by significantly different means”. Del resto, anche i calcolatori del tempo, impegnati in un semplice calcolo aritmetico, non avevano nessuna corrispondenza necessaria con la struttura degli elementi del cervello che eseguono quelle funzioni, e secondo Oettinger ci si sarebbe potuto aspettare che in futuro molte macchine avrebbero continuato ad avere solo una rassomiglianza funzionale con gli organismi viventi.⁴⁹

Costruire calcolatori poteva avere quindi scopi diversi: su un piano teorico, le simulazioni su calcolatore potevano essere strumenti per comprendere il funzionamento della mente umana; su un piano applicativo, si potevano imitare certe facoltà mentali per fini esclusivamente ingegneristici.

Chiudendo il suo discorso, Oettinger infatti affermava che:

In attempting to follow the example of the nature through the study and simulation of animal nervous system, we can expect to learn more about her ways,

⁴⁸ Ivi.: 242.

⁴⁹ Ibid.

and this increase knowledge may, in turn, benefit the computer art. But it can be expected that many computing and control problems will continue to be solved most easily and most economically by the use of structures unknown in living organisms.⁵⁰

Secondo Oettinger, concentrandosi sul piano prettamente teorico, i calcolatori potevano diventare strumenti per testare ipotesi riguardanti il funzionamento del cervello, come dei modelli logici per la neurologia e la psicologia, distinguendo due casi. Nel primo, se una teoria delle funzioni cerebrali era formulata in forma matematica, come quella di Mosteller e Bush sull'apprendimento condizionato, il computer poteva essere usato per risolvere semplici problemi di calcolo, come le equazioni differenziali, per ottenere valori numerici, e così via. Oppure, essa poteva essere, per così dire, in forma verbale, come quella di Hebb sull'apprendimento e la formazione di concetti. Nel 1948 Hebb aveva proposto l'ipotesi che le connessioni presenti nel cervello cambino continuamente via via che l'organismo apprende diversi compiti funzionali, e che questi cambiamenti creino insieme organizzati di cellule. In questo caso, il calcolatore digitale avrebbe potuto essere programmato per simulare la rete neurale con il suo ambiente, con lo scopo di testare la teoria di Hebb, che è proprio ciò di cui si occupò Rochester durante il simposio.

Nel suo discorso su *Simulation of Brain Action on Computers*, Rochester

⁵⁰ Ibid.

introduceva la metodologia ciclica nella simulazione delle reti neurali tramite calcolatori digitali:

If the model which represents the theory works as (Hebb) says it should, the experiment gives support to the theory. Of course it could not prove that the theory correctly represents living nets. On the other hand, if the model does not work as expected it might force modifications in the theory.⁵¹

La parte iniziale della simulazione consisteva infatti nella formulazione di un modello teorico circa il funzionamento del cervello, di cui il calcolatore avrebbe determinato le implicazioni, testato di dati, e infine approvato, disapprovato o modificato la teoria di partenza, ripetendo il circolo da capo, in un ciclo costante di auto-miglioramento.

Questa metodologia modellistica verrà ripresa da Newell e Simon qualche anno dopo al diverso livello che caratterizzò il loro approccio simulativo, quello della teoria dell'elaborazione dell'informazione (*Information Processing Theory*), la quale concepiva il soggetto umano come un elaboratore d'informazioni e guardava ai programmi per calcolatori come modelli per il funzionamento della mente, soffermandosi però sulla costruzione di modelli psicologici piuttosto che neurali dell'intelligenza. Dal loro punto di vista, una teoria del comportamento espressa come un modello di simulazione computerizzata era, in generale, la migliore alternativa sia per le descrizioni non solo matematiche e quantitative, come

⁵¹ Ivi.: 249.

quelle di Bush e Mosteller, ma anche per quelle verbali e qualitative, come quella di Hebb, che avevano il difetto, da una parte, di non descrivere adeguatamente un sistema di per sé piuttosto complesso e, dall'altra, la difficoltà di replicarne empiricamente i meccanismi in modo da far nascere dei fenomeni osservabili. Saranno gli stessi Newell e Simon, nel loro *Computers in Psychology*, a parlare dell'esperimento effettuato da Rochester durante il simposio, condotto su un IBM 704. La memoria del calcolatore venne divisa in 69 parti, rappresentanti un singolo neurone, e ognuno di essi era connesso con altri 10: il primo esperimento dimostrò che il postulato di Hebb non era sufficiente, perché i neuroni non sviluppavano spontaneamente degli insiemi cellulari, rendendo quindi necessarie delle modifiche alla teoria.⁵² Secondo Newell e Simon, dato un assioma matematico e vista la complessità del comportamento umano, vi erano più possibilità di costruire una teoria usando dei programmi computerizzati che con un tentativo diretto di formulazione matematica, considerando anche che i calcolatori, a differenza degli esseri umani, hanno il vantaggio non da poco di poter riportare i diversi passaggi del loro lavoro di elaborazione, rendendo l'analisi del problema ancora più abbordabile.⁵³ Per quanto riguarda la descrizione verbale della teoria, essa poteva essere sostituita con un programma computerizzato, o un modello che ne può portare alla luce inconsistenze e lacune, come dimostrato da Rochester. Si trattava,

⁵² Newell, Simon 1963: 396-397.

⁵³ Newell, Simon 1955: 22-23.

insomma, di stabilire le relazioni tra teoria e modello all'interno di una simulazione computerizzata simbolica, a partire da (pochi) fatti empirici e ipotesi teoriche circa il funzionamento della mente. In questo senso, formulare una teoria psicologica nei termini di un programma computerizzato corrispondeva all'affermare una teoria dell'elaborazione dell'informazione.⁵⁴

Al simposio, Rochester notò anche come la teoria di Hebb provasse a colmare la distanza tra psicologia e neurofisiologia.⁵⁵ Anche qui, la simulazione tramite calcolatori poteva contribuire allo scopo. Nella sua visione, la relazione tra le due discipline era sintetizzata dagli stati intermedi che intercorrevano tra di esse. Lo schema partiva con un Teorema A, un assunto riguardo il comportamento dei neuroni, che non contrastava con gli assiomi della neurofisiologia corrente (*established neurophysiology*). Questa asserzione, che poteva essere identificata con il postulato di Hebb, veniva allargata per includere il Teorema B, riconducibile a un livello meno microscopico, come quello dello sviluppo degli insiemi cellulari. Di nuovo, il Teorema B poteva essere allargato per includere qualcosa riguardante il comportamento, un Teorema C, e testato contro qualcosa che Rochester chiamava psicologia corrente (*established psychology*). Tramite la simulazione computerizzata si poteva testare la transizione da A e B, al fine di ridurre la distanza tra psicologia e

⁵⁴ Cordeschi 1997: 268.

⁵⁵ McCulloch et al. 1956: 243.

neurofisiologia nella teoria di Hebb. Ma come si è visto prima, fu proprio questa prima transizione a fallire nella simulazione, rendendo necessarie delle modifiche al postulato originale.

La distinzione tra diversi livelli di spiegazione o di studio (“psicologico” o “neurologico”) dell’intelligenza era già nell’aria, e si avviava a ispirare diversi programmi di ricerca, ma in quello specifico momento, evidentemente, non era avvertita come un motivo di contrapposizione tra i ricercatori. Qualche anno più avanti, Simon scriverà:

The use of computers to perform “humanoid” tasks falls into a number of distinguishable, though overlapping, categories. On the one hand, the goal may be to learn about human processes by simulating them; this has been the central motivation in simulating neural nets and a good part of the work on simulating human problem solving. On the other hand, the goal may be to find effective machine processes for accomplishing complex tasks-imitating the human processes only when this proves the most efficient way to do the job.⁵⁶

La teoria dell’elaborazione dell’informazione aveva l’enorme vantaggio di dire poco sui processi neurofisiologici e biochimici sottesi che si verificano nel sistema nervoso centrale e periferico.⁵⁷ L’esempio usato da Simon era quello di Mendel, il quale costruì la teoria dei geni dominanti e recessivi in base alla semplice osservazione e sperimentazione: fu solo molti anni dopo che l’osservazione delle strutture microscopiche delle cellule diede

⁵⁶ Simon 1961: 111.

⁵⁷ Ivi.: 113.

fondamento ai suoi studi. Sebbene, come già affermato da Oettinger, lo scopo dei due metodi era in entrambi i casi quello di spiegare il comportamento, la teoria dell'elaborazione dell'informazione approcciava questa spiegazione in più fasi:

They first reduce the complex behavior to symbol manipulating processes that have not, as yet, been observed directly in the human brain. The hope, of course, is that when we know enough 'about these processes, it will be possible to explain them at a still more fundamental level by reducing them to systems of neural events.⁵⁸

Così facendo, le teorie psicologiche avrebbero potuto essere dotate di una struttura gerarchica tipica di scienze quali la genetica e la biofisica, con al vertice la teoria dell'elaborazione dell'informazione, seguita dalle teorie neurologiche e da quelle biochimiche. La teoria dell'elaborazione dell'informazione si veniva quindi a configurare come un livello di spiegazione intermedio tra il microlivello della neurologia e quello macroscopico del comportamento manifesto: l'unico livello in cui era possibile una simulazione che provvedesse una verifica funzionale della teoria psicologica in assenza di dati empirici più solidi e strutturali, tramite la neonata programmazione euristica.

Durante la *Western Joint Computer Conference* del 1955, Pitts affermò che:

⁵⁸ Ibid.

The speakers this morning are all imitators in the sense that the poet in Aristotle "imitates" life. But, whereas Messrs. Farley, Clark, Selfridge, and Dinneen are imitating the nervous system, Mr. Newell prefers to imitate the hierarchy of final causes traditionally called the mind. It will come to the same thing in the end, no doubt.⁵⁹

Ma la separazione tra le due tendenze nella simulazione computerizzata era ben lungi dal sanarsi: era appena iniziata.

1.6 Calcolatori e flessibilità

Al simposio del 1955, le domande sollevate da Otto Schmitt nel suo discorso su *The Brain as a Different Computer* furono molto dibattute.

Come biologo egli constatava delle divergenze insanabili tra gli ordinari calcolatori digitali e il cervello biologico, utilizzando quindi un punto di vista diametralmente opposto a quello adottato da Oettinger:

I believe we have fallen into the trap of describing some brain functions in terms of present-day computer components and are then delighted to discover machine-like components in our description of brain function.⁶⁰

Per Schmitt, questi erano i classici errori in cui ci si imbatte quando si usa un'ipotesi camuffata da teoria per provare un'altra ipotesi.⁶¹ Pur ammettendo che le proprietà comportamentali dei neuroni possano corrispondere, solo in alcuni casi particolari, a quelle dei calcolatori digitali,

⁵⁹ Pitts 1955: 108.

⁶⁰ McCulloch et al. 1956: 244.

⁶¹ Ibid.

Schmitt sottolineava che questi ultimi avrebbero dovuto imitare la flessibilità nei ragionamenti solitamente dimostrati dagli esseri umani per poter essere dei buoni simulatori delle funzioni cerebrali. I calcolatori avrebbero quindi dovuto mostrare una sorta di “logica grigia”, non quella rigida, bivalente, o “bianca o nera” che li caratterizzava. Questo avrebbe permesso ai calcolatori di cogliere concetti astratti e poco definiti, così come la capacità di sfruttare il bagaglio di conoscenze spesso incomplete e conflittuali di cui generalmente dispongono gli esseri umani nella vita reale, e in base alle quali, nonostante tutto, sono comunque costretti a giungere a una decisione. Oettinger era perplesso circa l’utilità di una macchina dotata di una simile logica. Se una macchina in grado di operare secondo quella logica grigia poteva dare una simulazione realistica del comportamento cerebrale, allora essa avrebbe potuto rivelarsi di una certa utilità, mentre se la macchina era intesa come strumento per amplificare le capacità umane fisiche e mentali, allora vi erano già più incertezza circa il suo effettivo valore.⁶² Schmitt, in tutta risposta, giudicava che una siffatta macchina poteva essere utile a entrambi gli scopi, sia per fornire una realistica simulazione del comportamento del cervello (la forza del cervello umano risiedeva per lui in quella capacità di “improvvisazione”), sia per costruire macchine efficienti, come aiuto per i compiti che impiegano le capacità umani intellettuali, per i quali i calcolatori dell’epoca si rivelavano “irrimediabilmente ponderosi” e “stupidamente logici” nella loro necessità

⁶² Ivi.: 247.

di una programmazione precisa e nella loro mancanza d'immaginazione.⁶³

Nell'ultimo caso, Schmitt concludeva che:

It is necessary to abandon the idea of perfectly correct, uniformly logical solutions in any machine which is to arrive at generally appropriate quick solutions to complex problems when provided only with sketchy, conflicting, and partially inappropriate information and instructions.⁶⁴

Schmitt faceva l'esempio di un guidatore che può decidere di infrangere i limiti di velocità, data una situazione particolare. La decisione è ovviamente facile da prendere per una persona, ma sarebbe stata molto più difficile per un calcolatore. L'opportunità (e il modo) di includere la casualità in un calcolatore era un altro dei temi che saranno sollevati dalla Proposta di Dartmouth:

A fairly attractive and yet clearly incomplete conjecture is that the difference between creative thinking and unimaginative competent thinking lies in the injection of a some randomness [...] A problem whose solution requires originality could yield to a method of solution which involved randomness.⁶⁵

Un'altra delle questioni toccate da Schmitt riguardava la capacità di apprendimento nei calcolatori e negli organismi organici. Questi ultimi, grazie alla loro capacità di *problem-solving*, erano in grado di modificare la propria struttura e di conseguire dei successi e delle conoscenze poi

⁶³ Ibid.

⁶⁴ Ibid.

⁶⁵ McCarthy et al. 1955: 2, 8.

tramandate alla generazione successiva: una capacità preclusa ai calcolatori.

Di fronte a questa constatazione, Oettienger rispondeva che:

With computers, it seems to me that we are able in principle, by the use of appropriate programming or designing of structure, to build in one swoop the whole background of explicit existing knowledge. It seems wasteful to build a machine to repeat the trials and errors made by untold numbers of human brains in many thousand years.⁶⁶

Schmitt restava irremovibile:

Computers don't do any particular profound operations. Suppose that you wanted... you insert the knowledge of what 'democracy' is into a computer: would it not be impossible? It would certainly be very difficult...⁶⁷

Anche qui, come incarnare la capacità di apprendimento, i concetti astratti, e come riprodurre il background di conoscenze esplicite tramite i calcolatori, erano alcuni dei principali nuclei tematici della Proposta di Dartmouth, e sarebbero stati al centro delle ricerche sull'AI nei decenni a venire:

An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.⁶⁸

⁶⁶ McCulloch et al. 1956: 249.

⁶⁷ Ibid.

⁶⁸ McCarthy et al. 1955: 1.

1.7 Decisioni complesse

Un altro simposio, sempre sponsorizzato dal PGEC, si tenne il 22 marzo del 1956, pochi mesi prima la Conferenza di Dartmouth. Era uno dei primi incontri che aveva come argomento *The Impact of Computers on Science and Society*, e molti invitati provenivano non solo dal mondo delle scienze accademiche, ma principalmente dal governo e dall'industria.

Durante il suo intervento, Allen V. Astin, fisico americano e direttore del United States National Bureau of Standards dal 1951 al 1969, sottolineava come i cambiamenti sociali a quel tempo in corso fossero probabilmente imputabili a una maggiore influenza del governo sulla società: un processo risalente alla Prima guerra mondiale, ma accelerato durante e dopo la Seconda guerra mondiale. Il governo americano era diventato il principale finanziatore e utilizzatore delle scoperte in campo scientifico e tecnologico, il che rendeva il calcolatore digitale “a magnificent tool” che poteva “greatly assist government in its scientific research activities and in many of its other operations”.⁶⁹ Questi, tuttavia, presentavano delle pesanti limitazioni. Da una parte, vi erano problemi di natura scientifica e nell'elaborazione dei dati (*data-processing*) che non potevano essere risolti efficacemente da questi dispositivi. D'altra parte, un problema più prettamente di tipo culturale: i calcolatori erano ancora troppo

⁶⁹ Astin et al. 1956: 143.

“sottovalutati” nella percezione comune, tanto che “it takes considerable urgency and justification before one considers converting standard procedures to automatic methods”.⁷⁰ Si pensava ancora che i calcolatori fossero solamente delle grosse macchine di calcolo, utili per le applicazioni economico-finanziarie, ma molto meno in quelle aree, governative e industriali, in cui l’elaborazione di dati complessi e le procedure di ottimizzazione erano essenziali.

Al tempo, le principali tecniche di supporto in questi compiti provenivano dalla Ricerca Operativa (*Operational Research*), una branca della matematica che studia, su base quantitativa, i modelli concettuali dei processi decisionali connessi al funzionamento dei sistemi organizzati, i metodi per prevedere il comportamento di questi sistemi per individuarne le decisioni che ne ottimizzano le prestazioni, nonché gli strumenti per valutare a priori le conseguenze di determinate decisioni.⁷¹

Durante il simposio, è possibile notare una prima interazione tra OR e AI, ancora al suo stato “germinale”. Nel suo svolgersi veniva infatti usato esplicitamente il termine “intelligenza artificiale”, forse per la prima volta pubblicamente, appena qualche mese prima la Conferenza di Dartmouth. A usarla fu John Mauchly, uno dei costruttori dell’ENIAC, nel suo intervento in risposta a David Sayre, al tempo all’IBM e uno dei creatori del FORTRAN. La questione sollevata riguardava appunto l’uso dei calcolatori

⁷⁰ Ibid.

⁷¹ <http://www.treccani.it/enciclopedia/ricerca-operativa/>

nello studio delle procedure di *decision-making* per la risoluzione di problemi complessi e di pianificazione.

L'espressione "intelligenza artificiale" viene solitamente attribuita a McCarthy e compare nel documento di proposta della Conferenza di Dartmouth del 1955. Il nome di Sayre era presente nella lista delle persone che avrebbero potuto essere tra i partecipanti, in quanto interessati al programma di ricerca sull'AI.⁷² Al simposio, egli toccò infatti il tema delle macchine intelligenti, speculando sul modo di conferire a una macchina quella che lui chiamava "something that approaches intelligence":

I would envisage placing a machine in an environment which it can affect by its actions and from which the consequences of its actions are fed back to it. It would begin with a procedure that had been given to it, but it would at the same time execute neighboring procedures and test out by its interaction with its environment whether one of the neighboring procedures might not be more successful than the one it had executed; after enough favorable evidence it would adopt this procedure.⁷³

Una macchina del genere avrebbe avuto il merito di dimostrarsi abbastanza flessibile, e sarebbe stata dotata di quell'abilità di auto-miglioramento che caratterizzava la macchina "truly intelligent" che "will carry out activities which may best be described as self-improvement" cui si allude nella

⁷² McCarthy et al. 1955: 15.

⁷³ Astin et al. 1956: 157.

Proposta di Dartmouth.⁷⁴ Sayre, però, collegava esplicitamente una macchina del genere al *decision-making* tipico della OR, quando suggeriva che attività complesse, come quello della risoluzione dei problemi e della pianificazione, richiedevano “a rather different technique of machine use than we have yet developed”.⁷⁵ Stabilito che nessuna “procedura esatta” (*exact procedure*) era ancora stata sviluppata per risolvere questi problemi, la questione riguardava “how to cause a machine, which has been given a fairly exact procedure, itself to amplify and correct it, constantly producing better and better procedures”.⁷⁶ Queste procedure “abbastanza esatte” o “inesatte” erano considerate da Mauchly quale marchio distintivo dell’AI:

It is certainly true that many of us are interested in what has been given the name “artificial intelligence”. This is indeed a field in which a great deal is going to be done, and there will be much influence on the future applications if we are successful in some of the endeavors which (Sayre) described as coming under the “inexact” rules, procedures, and applications.⁷⁷

Il punto qui era la capacità dei calcolatori di prendere decisioni, simulando le procedure degli esseri umani nella risoluzione dei problemi, per assistere gli stessi nell’elaborazione d’informazioni complesse, nella pianificazione, e nel processo decisionale. Era inoltre necessario stabilire quale metodo sarebbe stato più efficace per il raggiungimento di questo scopo. Simon,

⁷⁴ McCarthy et al. 1955: 2.

⁷⁵ Astin et al. 1956: 149.

⁷⁶ Ibid.

⁷⁷ Ivi.: 155.

anni dopo, illustrerà i limiti della OR nell'affrontare queste situazioni complesse, dove l'informazione era in genere lacunosa, e le procedure non garantivano l'ottimizzazione del processo decisionale. Per Simon, il modello migliore da usare nello studio del processo decisionale non era l'economista onnisciente, l'*Homo oeconomicus*, che massimizzava le sue scelte come previsto dalla teoria dei giochi e assimilabile al *modus operandi* tipico della OR. Dotato di una razionalità ideale, si assumeva che l'uomo economico fosse sempre pienamente informato sull'insieme dei problemi e sull'ambiente, per quanto complessi potessero essere. Si trattava di un modello che poteva rivelarsi sicuramente utile nelle indagini, per studiare come gli individui e le organizzazioni sviluppano i propri obiettivi e metodi, descrivendo e giudicando diverse alternative, per infine riunire e utilizzare le informazioni raccolte per trarre conclusioni e decisioni. Ma la razionalità andava vista come una variabile, non come una costante: nessuno, che si trattasse di singoli individui o di entità più eterogenee, era in grado di soddisfare pienamente i requisiti richiesti dal modello classico del processo decisionale. Simon proponeva quindi un altro modello, più realistico: quello dell'"uomo amministrativo", dotato di una razionalità limitata, che lo rendeva incapace di massimizzare sempre le proprie scelte, rendendo necessaria l'introduzione di procedure decisionali "soddisfacenti", o euristiche. Da quel momento in avanti, le discipline scientifiche saranno solite usare uno dei due diversi modelli del processo decisionale sopracitati:

da una parte, l'OR, basata sulla comune programmazione lineare e sulle tecniche della teoria della probabilità e, dall'altra, l'AI, basata sulla neonata programmazione euristica. Scriveva Simon:

OR and AI each went their own way and seldom communicated until, less than a decade ago (*The future of information systems* è del 1997), expert systems that were developed in AI began to come to the attention of the public and of OR specialists. Now we have a second chance to bring these two disciplines and their tools together.⁷⁸

Quest'incapacità dell'OR nel trovare procedure di *problem-solving* più "umane" (cioè più vicine a come effettivamente avvengono nella realtà) fu la causa principale del suo primo "divorzio" con l'AI.⁷⁹

Alla vigilia della Conferenza di Dartmouth e sullo sfondo della diffusione dei primi calcolatori digitali erano già state sollevate molte delle questioni riguardanti l'AI che avrebbero influenzato sia il futuro della ricerca sia quello delle controversie negli anni a venire.

Come si è visto, il risultato di queste trasformazioni fu la nascita della programmazione euristica, che puntava alla ricerca di soluzioni soddisfacenti anche se non ottimali, a partire da informazioni incomplete e da dati non sempre empiricamente confutabili. Il *Logic Theorist* è solitamente considerato il primo programma euristico della storia ma, anche se giocò un ruolo importante nella Conferenza di Dartmouth, programmi

⁷⁸ Simon 1997: 6.

⁷⁹ Cordeschi 1997: 274-276.

precedenti e differenti per complessità, da Oettinger a Samuel, includevano procedure definibili come euristiche. Questi sviluppi furono anche tra le principali cause del prevalere della simulazione dei comportamenti umani tramite calcolatore a discapito del *brain modelling* e delle reti neurali tipiche della cibernetica. Queste ultime iniziarono a essere presto considerate come uno stile di elaborazione eccessivamente “cerebrale”, in particolare quando l’AI, come nuova scienza del pensiero, assunse lo status di livello di spiegazione comportamentale autonomo del sistema nervoso. Due anni dopo Dartmouth, al *Teddington Symposium* del 1958, questa distinzione tra imitazione del cervello e imitazione del pensiero sarà stabilita una volta per tutte da Minsky:

Most of the systems discussed up to this point may be regarded as dealing with rather clear-cut syntactic processes involving the manipulation of symbolic expressions. But other kinds of adaptive machines have been studied for which there seem to be no such natural linguistic descriptions.⁸⁰

Si trattava, appunto, delle reti neurali, composte da ampi assemblaggi di componenti interconnessi, molti dei quali potevano operare nello stesso momento. Ma, in assenza di una rappresentazione “fisica” dei concetti di ordine superiore (*higher order concepts*), da queste reti auto-organizzate era possibile ottenere solo un apprendimento elementare e altri tipi di

⁸⁰ Minsky 1959: 24.

comportamenti adattivi.⁸¹ Hebb fu il primo a sviluppare delle teorie che andavano in quella direzione, ma la conoscenza del cervello umano era ancora troppo limitata per conseguire dei successi significativi.

Concludeva Minsky:

This is one reason why the author feels that, even for those whose central interest is in unraveling the mysteries of the brain, it might be well to devote a major share of the effort, at the present time, to the understanding and development of the kind of heuristic considerations that some of us call “artificial intelligence”.⁸²

⁸¹ Ivi.: 25

⁸² Ivi.: 26.

2. Gli Automata Studies

Nel periodo compreso tra il 1945 e il 1956, con l'affermarsi del concetto di elaborazione simbolica, l'AI si andava delineando come campo di ricerca indipendente, e furono tre gli eventi cruciali che permisero il raggiungimento di questo traguardo.

Il primo di essi è stato il tema centrale dello scorso capitolo: l'AI doveva legarsi ai calcolatori. Le prime indagini sulle macchine intelligenti si concentrarono principalmente sulla teoria della retroazione e nel replicare il funzionamento del cervello attraverso la riproduzione di reti neurali artificiali. I pionieri dell'AI iniziarono lentamente (e quasi inconsciamente) a rigettare questo approccio quando si accorsero, tramite il lavoro e lo studio diretto sui primi calcolatori digitali, che i processi del pensiero umano potevano essere replicati più efficacemente attraverso questi dispositivi, al tempo nel pieno della loro diffusione.

Seconda cosa, doveva svilupparsi una comunità scientifica che si rispecchiasse in questa nuova prospettiva. Minsky disse che l'avvento dei calcolatori negli anni Cinquanta “made it practical to be much more ambitious”, ma mancava ancora una qualche forma di riconoscimento, di fondazione ufficiale (o ufficioso). In questo periodo le differenze fondamentali tra cervello umano, sistemi a retroazione e calcolatori digitali non erano ancora del tutto chiare, tanto che la prima generazione di ricercatori dell'AI, cui spettò il compito di rimarcare queste differenze,

aveva comunque dei pesanti debiti nei confronti della cibernetica. Questo primo gruppo comprendeva Marvin Minsky, John McCarthy, Allen Newell, Herbert Simon e i loro studenti.

Terza cosa, e forse più importante: tali personalità dovevano incontrarsi e confrontarsi. Questo processo culminò nel 1956 con la Conferenza di Dartmouth, durante la quale venne presentato e discusso il primo programma euristico della storia: il *Logic Theorist*.

2.1 Claude Shannon

Claude Elwood Shannon viene generalmente considerato il fondatore della teoria dell'informazione di scuola americana. Il cuore della sua teoria risiedeva in una nuova concettualizzazione del termine "informazione". Per rendere la teoria della comunicazione una disciplina scientifica, Shannon aveva bisogno di trasformare l'informazione in un parametro fisico quantificabile, e ciò fu reso possibile tramite la distinzione semantica tra informazione e significato (*meaning*). Con "significato" Shannon intendeva il contenuto specifico di un particolare messaggio, mentre l'"informazione" era il numero di messaggi che potevano essere trasmessi su un canale: maggiore era il numero di messaggi possibili, maggiori sarebbero state le informazioni trasmesse, perché si sarebbe potuti risalire al contenuto originale scartando più alternative.⁸³

Sebbene gli studi sulla teoria dell'informazione occupassero gran parte

⁸³ Aspray 1985: 119-120.

della sua vita, Shannon ebbe sempre un grande interesse per la cibernetica e la teoria degli automi. Nato a Petoskey nel 1916, egli mostrò già da bambino una certa predisposizione per la scienza, in particolare per tutto ciò che riguardava gli apparecchi meccanici, tanto da costruire, prima della maggiore età e con materiali di fortuna, una nave giocattolo telecomandata e un telegrafo.⁸⁴ Cresciuto con i miti di Edison, Newton, Darwin ed Einstein, nel 1938, all'età di ventidue anni, venne assunto come assistente da Vannevar Bush.

Nel 1931, Bush aveva progettato l'analizzatore differenziale, in grado non solo di sommare e moltiplicare numeri, ma anche di risolvere equazioni differenziali tramite l'utilizzo di alberi rotanti come dispositivi per l'integrazione, rendendola una delle prime macchine di calcolo a essere impiegate operativamente.⁸⁵ La macchina di Bush era di fatto un calcolatore analogico, dal momento che la posizione angolare di un albero poteva assumere qualunque valore, e la precisione della macchina era limitata solo da quanto correttamente si potesse calcolare la loro posizione.

Ma anche se il cuore della macchina era meccanico e analogico, essa era controllata da un circuito digitale che faceva uso di relè elettromeccanici. Questo fatto diede a Shannon una forte intuizione: perché non rendere la macchina meno complessa, lasciando l'elaborazione ai soli relè? Queste considerazioni lo portarono a dimostrare che si sarebbe potuto costruire,

⁸⁴ Sloane, Wyner 1993: xi.

⁸⁵ Hodges 2006: 207.

usando solo degli interruttori interconnessi che potevano essere accesi o spenti, una macchina a calcolo digitale capace di eseguire ogni possibile operazione numerica. Gli interruttori avevano il vantaggio di operare molto più velocemente rispetto alle ruote dentate precedentemente usate nelle macchine digitali. Ma dal momento che essi potevano assumere solo due posizioni (accesi o spenti), erano possibili solo due valori (0 e 1) per la rappresentazione digitale dei numeri, ed è per questo che inizialmente i calcolatori usavano l'aritmetica binaria. Fu da questo momento che relè e interruttori presero a sostituire gli ingranaggi nelle operazioni di calcolo meccanico.⁸⁶

Come visto nel precedente capitolo, Shannon iniziò presto a interessarsi ai calcolatori digitali, scrivendo articoli come *Programming a Computer for Playing Chess* e instaurando nel frattempo una corrispondenza con Wiener (che era stato suo professore al MIT) sull'argomento.⁸⁷ Nei primi anni Cinquanta partecipò a tre conferenze Macy, dove presentò i risultati delle sue ricerche riguardanti la teoria dell'informazione e quella degli automi; qui ebbe modo di conoscere McCulloch, Pitts, Walter, Bigelow e Mackay, uno dei fondatori della scuola inglese della teoria dell'informazione che già in precedenza aveva compiuto delle ricerche sugli automi.⁸⁸ Dopo aver partecipato alla conferenza Macy del 1949 in qualità di oratore, McCulloch

⁸⁶ Crevier 1993: 11-12

⁸⁷ Kline 2011: 6.

⁸⁸ Heims 1994: 11, 75-76.

invitò Shannon a presenziare alle due successive come ospite. Nei loro scambi epistolari, Shannon scriveva che:

There is a considerable activity here (ai Bell Labs) in information theory, computing machines and other ‘Cybernetic’ questions, and a fresh viewpoint is stimulating.⁸⁹

La conferenza Macy del 1950 aveva come tema il linguaggio, e la teoria dell’informazione di Shannon si occupava, tra le altre cose, della sua struttura statistica. Shannon fu accolto piuttosto positivamente dal gruppo dei cibernetici. Solo Pitts lo incalzò con domande di natura tecnica, mentre Wiener confrontò i risultati di Shannon con il suo metodo di previsione delle serie temporali.

Sempre nel 1950 fu la volta dell’articolo *Presentation of a Maze-Solving Machine*, in cui Shannon illustrava il funzionamento della sua ultima creazione: si trattava di *Theseus*, un topo meccanico comandato da relè, la cui peculiarità era quella di riuscire a orientarsi nel labirinto in cui era stato posto. In *Theseus*, sia il “cervello” che i “muscoli” erano separati dal topo stesso e collocati sotto il labirinto. Il cervello era un circuito di circa 100 relè, e i muscoli costituiti da una coppia di motori guidati da un elettromagnete che muoveva il topo attraverso il labirinto. Quest’ultimo, costituito da 25 caselle, poteva essere modificato a piacimento: il topo

⁸⁹ Tratto dai *Warren McCulloch Papers* 2 giugno 1949; la corrispondenza private di McCulloch è custodita presso l’American Philosophical Society Library.

avrebbe cercato, al suo interno, il posto arbitrariamente prefissato come obiettivo da raggiungere, riuscendo a orientarsi. Anche il topo poteva essere collocato in qualunque punto del labirinto; vagando per esso alla ricerca di un posto già conosciuto, avrebbe poi proceduto verso la meta, immagazzinando le nuove conoscenze nella sua memoria.⁹⁰ Per via di queste capacità, *Theseus* viene generalmente considerato uno dei primi dispositivi capaci di apprendimento.⁹¹

Sebbene il suo interesse primario fosse la teoria dell'informazione (e tale sarebbe rimasto per tutta la vita), Shannon non abbandonò mai del tutto la cibernetica, tanto da scrivere, nel 1957 e su raccomandazione di Wiener, un lungo articolo per la prestigiosa *Encyclopaedia Britannica* avente come oggetto proprio la cibernetica. Nell'articolo, veniva descritta usando la stessa definizione coniata da Wiener, ovvero come la “science of control and communication processes in both animal and machines”. Riguardo i confini tra cibernetica e teoria dell'informazione, Shannon affermava che la cibernetica si “sovrappone” (*overlaps*) a campi quali la neurologia, la scienza dei calcolatori e la teoria dell'informazione. Per Shannon, la cibernetica era “universale”, in quanto cercava di trovare le caratteristiche comuni a queste diverse discipline.⁹²

L'interesse di Shannon nei confronti della cibernetica aiuta a capire meglio i motivi che lo portarono, nei primi anni Cinquanta del secolo scorso, a co-

⁹⁰ Shannon 1951: 681.

⁹¹ Sloane, Wyner 1993.: xiv

⁹² Shannon 1957: 916.

editare con McCarthy gli *Automata Studies*: evento fondamentale che convincerà quest'ultimo a organizzare la Conferenza di Dartmouth del 1956.

2.2 John McCarthy

McCarthy nacque a Boston nel 1927, pochi anni prima l'inizio della Grande depressione. Come molti *enfant prodige*, la sua formazione fu in buona parte da autodidatta. A causa di una malattia, fu costretto a iniziare la scuola con un anno di ritardo, riuscendo però a diplomarsi due anni prima del previsto. Ancora adolescente, sviluppò un forte interesse per la matematica, tanto da decidere di iscriversi al Caltech (California Institute of Technology). A 15 anni comprò il libro di testo che l'università utilizzava per gli studenti di matematica del primo anno, apprendendo praticamente da solo tutte le nozioni in esso contenute. Nella domanda d'ammissione per la Caltech (l'unica università presso cui inviò la richiesta) scrisse una sola cosa: "I intend to be a professor of mathematics".⁹³ Prima dell'inizio delle lezioni, si imbatté in uno dei professori del primo anno, ponendogli una domanda molto articolata riguardante uno dei problemi affrontati nel libro di testo, e mostrando alcune possibili soluzioni su cui stava lavorando. Il risultato fu che né il professore del primo anno né quello del secondo lo vollero nella propria classe: semplicemente, McCarthy era troppo preparato, e loro non avevano nulla da insegnargli. Alla fine, all'età di appena 16 anni,

⁹³ Hayes, Morgenstern 2007: 94.

McCarthy iniziò a frequentare direttamente i corsi dell'ultimo anno. Dopo aver conseguito il *Bachelor of Science*, iniziò il suo dottorato, sempre al Caltech.

Per McCarthy, il fermento intellettuale dell'epoca era tangibile.

L'entusiasmo generato dall'utilizzo di primi calcolatori per decodificare i messaggi criptati dell'Asse era ancora molto alto. Inoltre, come si è visto, vi era un crescente interesse nei confronti della mente e dei processi sottesi al suo funzionamento.

Nel settembre 1948, McCarthy partecipò allo *Hixon Symposium on Cerebral Mechanisms in Behavior*, una conferenza che riuniva le personalità di spicco delle diverse aree di ricerca legate alle scienze cognitive, tra cui Turing, Shannon, e lo psicologo Karl Lashley. Ascoltando gli interventi in cui si mettevano a confronto cervello e calcolatori, McCarthy ebbe come un'illuminazione. Da quel momento in poi, il suo principale interesse sarebbe stato la creazione di macchine in grado di pensare come un essere umano. Alcuni degli esempi che in seguito avrebbe usato nei suoi articoli, come quello della scimmia e delle banane, provenivano proprio dallo *Hixon Symposium*. Karl Lashley descrisse un esperimento comportamentale in cui una scimmia, spostando una scatola e arrampicandosi su di essa, riusciva a raggiungere un casco di banane posizionato in alto: in futuro, McCarthy avrebbe spesso preso in prestito questo episodio utilizzandolo come esempio per descrivere il livello

basilare di un ragionamento intelligente.

Fu durante questo congresso che McCarthy cominciò a prendere in considerazione la possibilità di costruire dei calcolatori intelligenti, ma solo su un piano puramente ipotetico poiché, in quel momento della sua carriera, non aveva la possibilità di utilizzarne nessuno. Teorizzò invece il funzionamento di due automi a stati finiti in grado di interagire tra loro, in cui uno rappresentava il cervello umano, l'altro l'ambiente. Quando arrivò a Princeton nel 1949, parlò a John von Neumann delle sue idee; von Neumann si dimostrò molto interessato all'argomento e gli consigliò di scrivere qualcosa a riguardo.⁹⁴ Ma McCarthy, dopo ulteriori riflessioni, iniziò presto a provare insoddisfazione nei confronti della sua idea originale. Bisognava trovare un modo per rappresentare il bagaglio di conoscenze depositate nel cervello umano; una coppia di automi a stati finiti, pur potendo rappresentare con precisione il comportamento di un cervello umano che interagisce con l'ambiente, erano del tutto inadeguati allo scopo. Seppur non utilizzasse ancora questa definizione, McCarthy aveva già individuato la necessità di una rappresentazione adeguata della conoscenza. Vent'anni dopo, formalizzerà questa intuizione operando la distinzione tra una rappresentazione metafisicamente adeguata e una rappresentazione epistemologicamente adeguata:

⁹⁴ McCorduck 1987 :87

A representation is called metaphysically adequate if the world could have that form without contradicting the facts of the aspect of reality that interests us. [...]

A representation is called epistemologically adequate for a person or machine if it can be used practically to express the facts that one actually has about the aspect of the world.⁹⁵

McCarthy terminò il suo dottorato di ricerca in soli due anni, trasferendosi successivamente a Princeton, dove divenne docente. Fu proprio durante questo periodo che, oltre von Neumann, conobbe anche Marvin Minsky, il quale aveva lì iniziato i suoi studi universitari proprio mentre McCarthy subentrava come insegnante. McCarthy e Minsky scoprirono presto di condividere la volontà di costruire macchine intelligenti; i due avrebbero collaborato su molti progetti nel corso del decennio successivo.⁹⁶

McCarthy trascorse l'estate del 1952 lavorando con Shannon ai Bell Labs. Qui gli propose l'idea di pubblicare insieme una raccolta di articoli aventi tema le macchine intelligenti. La corrispondenza tra i due riguardante gli *Automata Studies* può dare un'idea delle tensioni che si erano prodotte, ma mostra anche come Shannon possedesse una visione più ampia sull'argomento rispetto al più giovane McCarthy.⁹⁷

⁹⁵ McCarthy, Hayes 1969: 9.

⁹⁶ Hayes, Morgenstern 2007: 93-94.

⁹⁷ McCorduck 1987.

2.3 Il reclutamento per gli *Automata Studies*

A partire dal 1953, Shannon e McCarthy iniziarono a invitare diversi ricercatori che parteciparono alle conferenze Macy affinché contribuissero al volume: tra di essi spiccavano i nomi di Wiener, MacKay, Ashby, Walter e soprattutto von Neumann, che in quegli anni aveva iniziato a interessarsi alla teoria degli automi.

Agli inizi della sua carriera, tra il 1919 e il 1923, il neurofisiologo Warren McCulloch stava lavorando a un problema di logica filosofica: la creazione di un sistema formale che spiegasse l'utilizzo dei verbi transitivi.

Occupandosi di questo lavoro, finì con l'interessarsi anche a un problema a esso collaterale, che coinvolgeva la logica delle relazioni. In realtà McCulloch aveva in mente una teoria psicologica, e non filosofica, della logica relazionale. Un filosofo avrebbe probabilmente tentato di costruire un sistema formale che rispecchiasse l'utilizzo tipico della logica delle relazioni, ma le intenzioni di McCulloch erano diverse: egli mirava infatti a sviluppare una teoria che ne spiegasse le fondamenta psicologiche, non solo la struttura formale. Lo scopo di McCulloch era quello di poter teorizzare un evento psichico, uno "psicone" (*psychon*) avente le seguenti proprietà: esso doveva essere talmente semplice da poter dire che "it either happened or else it did not happen"; doveva verificarsi solo nel caso in cui fosse avvenuta la sua causa diretta, in modo da poter stabilire un antecedente temporale; questo evento psichico avrebbe potuto essere messo in relazione

con gli altri che ne sarebbero seguiti; infine, questi *psychon* avrebbero potuto essere legati tra di loro per produrre proposizioni più complesse rispetto alle antecedenti. Nel 1929, McCulloch decise di assimilare la meccanica degli psiconi a quella del tutto-o-niente tipica degli impulsi neuronali⁹⁸:

I began to try to formulate a proper calculus for these events by subscripting symbols for propositions (connected by implications) with the time of occurrence of the impulse in each neuron.⁹⁹

A ogni proposizione corrispondeva una linea di trasmissione neurale: a ogni proposizione vera corrispondeva una linea di trasmissione nello stato “uno”, mentre a una proposizione falsa lo stato “zero”. Così facendo i neuroni teorici diventavano il corrispettivo dei connettivi proposizionali, come le congiunzioni “e”, “oppure”, “né” ecc. Considerando ad esempio un neurone la cui soglia sia pari a 2, esso emetterà un impulso in uscita soltanto se entrambe le linee d'ingresso si trovano nello stato “uno”, cioè se sono entrambe vere, svolgendo le stesse funzioni di un connettivo. Agendo sulle due proposizioni, il neurone formerà con esse un composto che risulterà vero se e solo se entrambe le proposizioni di partenza sono vere.

Ma nella formulazione della sua psicologia delle proposizioni, McCulloch fu rallentato da molte difficoltà di natura tecnica. Fu l'incontro con il logico

⁹⁸ Secondo questa teoria, se la somma degli impulsi facilitatori in ingresso di un neurone è maggiore o uguale al valore della soglia, allora esso emetterà un impulso in uscita; viceversa, se ciò non si verifica, non vi sarà alcuna uscita.

⁹⁹ McCulloch 1960: 12.

Walter Pitts, che era in grado di fornire la preparazione matematica necessaria per la risoluzione di quei problemi, a sbloccare la situazione. Il risultato di questa collaborazione fu l'articolo del 1943 *A Logical Calculus of the Ideas Immanent in Nervous Activity*.

Usando come assiomi generali le regole che McCulloch aveva stabilito per i suoi psiconi e come quadro logico di riferimento un amalgama del calcolo logico di Carnap e dei *Principia Mathematica* di Russell e Whitehead, McCulloch e Pitts presentarono un modello logico di reti di neuroni funzionalmente simili alle macchine di calcolo di Turing. Come von Neumann sottolineò nel suo *General and Logical Theory of Automata*, l'essenza del contributo di McCulloch e Pitts stava nell'aver mostrato come qualsiasi operazione del cervello esprimibile, in modo chiaro e inequivocabile, con un numero finito di stadi, poteva essere ugualmente descritta tramite una delle loro reti neurali formali. Questo concetto di computabilità si rifaceva dichiaratamente a quello di Turing: tutte le computazioni realizzabili con le reti neuronali di McCulloch e Pitts erano, di fatto, coestensive della computabilità di Turing. Instaurare questa relazione era del resto uno degli obiettivi manifesti degli autori; già nel 1945 essi capirono che le reti neurali, quando fornite di un nastro infinito analogo a quello di Turing, erano equivalenti a quelle stesse macchine. Così come le macchine di Turing fornirono una teorizzazione astratta del pensare nel mondo meccanico, usando come modello di riferimento il pensiero

umano, adesso le reti neurali di McCulloch e Pitts ne fornivano una per il mondo biologico, usando come modello di riferimento una macchina: il risultato fu la nascita di una teoria unificata del pensiero capace di infrangere le barriere tra le teorie fisiche e quelle biologiche. Ma l'articolo non si limitava a sottolineare la somiglianza funzionale tra cervello umano e dispositivi di calcolo; fornì anche un nuovo modo di concepire il cervello come una macchina, utilizzando metodi di ricerca molto più precisi rispetto a quelli del passato, attraverso la formalizzazione matematica.¹⁰⁰

Il lavoro di McCulloch e Pitts va ovviamente contestualizzato: oggi sappiamo che le decisioni coscienti riconducibili alle preposizioni logiche avvengono a un livello molto più alto rispetto a quello del singolo neurone, e che probabilmente coinvolgono milioni di cellule cerebrali. Nonostante tutto, il loro lavoro fu senza dubbio seminale, perché riuscì a spiegare, tramite l'analisi matematica, come le cellule interconnesse, trasmettendo o meno gli impulsi, riescano a compiere delle operazioni logiche – come, cioè, il cervello, inteso come meccanismo, possa elaborare le informazioni ricevute.

Ma il loro parallelismo tra la capacità di elaborazione delle reti neurali artificiali e le macchine di Turing diede origine, tuttavia, alla falsa certezza che il cervello umano operasse esattamente, anche a livello biologico, come un calcolatore digitale, il quale perdeva così la semplice funzione di modello logico e teorico. Ci vorranno molti anni prima che le ricerche

¹⁰⁰ Aspray 1985: 129-130.

sull'AI si allontanassero da questo vicolo cieco. L'articolo di McCulloch e Pitts, all'inizio passato in secondo piano, fu reso popolare da von Neumann, diventando di conseguenza un punto di riferimento per tutti gli studiosi dell'intelligenza meccanica negli anni a venire. Era questa la realtà con cui McCarthy dovette scontrarsi durante la stesura degli *Automata Studies*: cosa che, come si vedrà, gli diede non pochi grattacapi e insoddisfazioni.

Tornando a von Neumann, è possibile dire che, mentre lo scopo di Wiener era quello di sviluppare una scienza interdisciplinare intorno al concetto di retroazione negativa (la cibernetica), il suo era quello di unificare il lavoro di Shannon, Turing, McCulloch e Pitts in un'unica teoria generale degli automi. Era, insomma, una delle figure con cui ci si doveva necessariamente confrontare se si voleva entrare nel dibattito scientifico del tempo, e non vi era nulla di strano che fosse uno dei candidati più probabili per gli *Automata Studies*, nonostante la visione di McCarthy sull'argomento fosse molto diversa dalla sua.

Shannon e McCarthy invitarono anche altri ricercatori di spicco impegnati nella teoria degli automi e delle reti neurali, incluso lo stesso Turing, che Shannon, come già accennato, ebbe modo di conoscere durante la guerra e che successivamente visitò in Inghilterra.¹⁰¹

Un altro candidato era Stephen Kleene, della University of Wisconsin-Madison, il quale aveva scritto un articolo sulle reti neurali dal titolo *Representation of Events in Nerve Nets and Finite Automata*: nato come

¹⁰¹ Hodges 2006: 461.

rapporto per la Rand Corporation nel dicembre 1951, esso sarebbe stato poi pubblicato all'interno degli *Automata Studies*. Klenne partiva dall'analisi delle reti neurali teorizzate da McCulloch e Pitts per dimostrare che esse erano capaci di risolvere a rigore solo una sottoclasse di problemi computabili da una macchina di Turing, cioè i cosiddetti eventi regolari (*regular events*): questi erano descrizioni costituite da una successione di simboli ottenute utilizzando i tre simboli logici primitivi di unione, prodotto di insiemi e iterazione.¹⁰²

Shannon e McCarthy contattarono anche Albert Uttley, neuropsicologo che in quel periodo stava lavorando a una rete neurale per il riconoscimento delle forme per l'esercito britannico, e Minsky, il quale accettò di contribuire al volume esponendo le proprie ricerche sulle reti neurali (che era anche l'argomento su cui si sarebbe basata la sua tesi di dottorato a Princeton). La scelta di includere Minsky (in quel periodo ancora studente) non era dettata solo dall'amicizia che lo legava a McCarthy. Le ricerche di Minsky erano infatti tenute in massima considerazione da Wiener, tanto da fargli scrivere in una lettera di raccomandazione del 1954 che sarebbe andato "considerably further than the best work to date of such a man as Ashby and Gray Walter".¹⁰³

MacKay, Ashby, Walter e Uttley erano membri del Ratio Club, un gruppo informale di scienziati che rappresentava gran parte del nucleo del primo

¹⁰² L'articolo, in realtà, presentava solo una prima esposizione di questo teorema, che successivamente sarebbe divenuto popolare come Teorema di Kleene.

¹⁰³ La lettera, indirizzata C. Brinton, è custodita presso il MIT.

movimento cibernetico britannico. Fondato da John Bates, neurologo del National Hospital, l'unico criterio di ammissione al Ratio Club era che solo "those who had Wiener's ideas before Wiener's book appeared" potevano farne parte. Non si trattava affatto di un gruppo di cibernetici dell'ultima ora: molti di loro si occupavano di questioni quali la riproducibilità dell'intelligenza e l'elaborazione dell'informazione nel cervello e nelle macchine già anni prima che Wiener coniasse il termine della nuova disciplina.¹⁰⁴ Il gruppo ebbe il suo primo incontro in occasione della visita di McCulloch in Inghilterra nel 1949. Successivamente vi entrò a far parte anche Turing, il quale vi presentò due relazioni. La prima, esposta nel 1950, era intitolata *Educating a Digital Computer*, e si soffermava su come integrare nelle macchine dei comportamenti adattivi in modo che imparassero e trasmettessero le informazioni ottenute alla generazione successiva, in quello che si potrebbe definire una sorta di evolucionismo artificiale.

La seconda relazione, *The Chemical Origin of Biological Form*, si concentrava su questioni di carattere morfogenetico, mostrando come modelli e forme potessero emergere da sistemi di reazione-diffusione se questi erano parametrizzati in modo appropriato (un ruolo che Turing ipotizzava potesse essere assunto dai geni). Oltre per l'aver fornito nuove direzioni alla biologia teorica, questo lavoro è pioneristico anche per l'uso estremamente originale che faceva dei calcolatori. Turing utilizzò il

¹⁰⁴ Husbands, Holland 2008: 90.

calcolatore digitale MkI della University of Manchester per risolvere numericamente alcune equazioni di reazione-diffusione non-lineari: uno sviluppo che si sarebbe rivelato estremamente influente negli anni successivi¹⁰⁵. E' proprio in questo periodo che Turing ricevette l'invito di Shannon e McCarthy. Egli, tuttavia, declinò l'offerta scusandosi: "I have been working on something quite different for the last two years, though I expect to get back to cybernetics very shortly"¹⁰⁶.

2.4 Una falsa partenza

McCarthy giocò un ruolo fondamentale all'inizio del progetto, assumendo praticamente il ruolo di *talent scout*, nonostante gli invitati godessero spesso di maggiore fama e prestigio rispetto a lui. Malgrado la sua giovane età, McCarthy non esitava a criticare, spesso duramente, gli articoli che non rientravano nella sua visione dell'argomento. Ad esempio, l'articolo di MacKay, *The Epistemological Problem for Automata*, discuteva su come un automa avrebbe potuto rappresentare simbolicamente l'ambiente esterno in maniera auto-organizzata e adattiva, proponendo un circuito analogico-elettrico per illustrare il principio di retroazione negativa su cui si basava la sua idea. McCarthy scrisse a Shannon:

¹⁰⁵ Turing 1952: 65; Husbands, Holland 2008: 106.

¹⁰⁶ Si riferiva, appunto, agli studi sulla morfogenesi biologica.

I think we should include it (l'articolo) with whatever changes he wishes to make. I don't think that the eventual solution to this problem will be an extension of his ideas, but I have no very precise grounds for this opinion.¹⁰⁷

Dopo aver criticato MacKay per le sue assunzioni circa le trasformazioni sensoriali, basate sul modello delle reti neurali di McCulloch-Pitts, McCarthy concludeva:

Nevertheless, I think the paper has a number of good ideas and also represents a point of view which cannot be dismissed as unfruitful this early in the game.

McCarthy ebbe più problemi con Ross Ashby, collega di MacKay. Ashby era l'autore di uno dei libri che influenzò maggiormente la storia della prima cibernetica: *Design for a Brain*, pubblicato nel 1952. Tra i suoi contributi scritti e la sua partecipazione attiva presso la comunità scientifica, Ashby è oggi considerato uno dei pionieri, se non il cofondatore, della cibernetica.¹⁰⁸

Ashby è famoso soprattutto per il suo omeostato, una delle prime macchine cibernetiche della storia. Completato nel 1948, esso consisteva in un sistema di quattro dispositivi elettromeccanici (il cui stato poteva essere variato tramite opportuni commutatori, che simulavano l'ambiente esterno) collegati elettricamente tra loro in modo che lo stato di ciascuno di essi

¹⁰⁷ Questa lettera (risalente al 5 aprile 1954), insieme alle altre riguardanti la stesura degli *Automata Studies* che verranno citate in seguito, è tratta dalla corrispondenza privata tra Shannon e McCarthy, custodita presso la Library of Congress, a Washington DC. Trattandosi di materiale che non ho avuto modo di visionare personalmente, e per dare comunque un'idea della successione temporale degli eventi, nelle prossime note sarà riportata solamente la data cui risale la lettera citata.

¹⁰⁸ Asaro 2008: 150

venisse a dipendere anche da quelli assunti dagli altri tre; se lo stato di uno dei dispositivi veniva allontanato dalla situazione di equilibrio prefissata, la variazione casuale della posizione dei quattro commutatori produceva una variazione della corrente che riportava ciascuno di questi in posizione di equilibrio, ottenendo però una configurazione complessiva diversa da quella iniziale.¹⁰⁹

Si trattava insomma di un sistema “ultrastabile”, in grado non solo di auto-correggere il proprio comportamento (come i sistemi dotati di retroazione negativa), ma anche di auto-organizzarsi, cioè di modificare la propria organizzazione interna al fine di reagire in modo appropriato ai cambiamenti esterni. Ciò comportava che la macchina, come rivelava il suo comportamento, era dotata di uno “scopo” nel senso cibernetic del termine; scopo che poteva anche non coincidere con quelli del progettista (una differenza che nella filosofia di Ashby si rivelerà fondamentale).¹¹⁰

Vista la sua fama, è probabile che Ashby interpretò l’invito di Shannon e McCarthy come un’accettazione automatica del suo articolo *Design for an Intelligence Amplifier*. Secondo Ashby, l’intelligenza consisteva nella capacità di operare selezioni appropriate:

It has often been remarked that any random sequence, if long enough, will contain all the answers. Nothing prevents a child from doodling “ $\cos^2 x \text{ þ } \sin^2 x \text{ } \frac{1}{4} 1$ ”.

[...] Doodling, then, or any other random activity, is capable of producing all that

¹⁰⁹ <http://www.treccani.it/vocabolario/omeostato/>

¹¹⁰ Asaro 2008: 160.

is required. What spoils the child's claim to be a mathematician is that he will doodle, with equal readiness, such forms as “ $\cos^2 x \text{ p } \sin^2 x \frac{1}{4} 2$ ” or “(ci)xsi-nx1” or any other variation. After the child has had some mathematical experience he will stop producing these other variations. He becomes not more, but less productive: he becomes selective.¹¹¹

Questa definizione rappresentava un radicale cambio di prospettiva rispetto alla concezione classica dell'intelligenza meccanica, intesa come la capacità di produrre risposte adeguate tramite la progettazione. L'intelligenza veniva ora intesa come l'unione tra la capacità di produrre un gran numero di alternative senza senso e quella di eliminare, tramite selezione appropriata, le possibilità non corrette: un processo in due fasi.

Nell'articolo, Ashby estendeva questa concezione al suo omeostato, che selezionava casualmente i propri stati interni fino a trovare quello che avrebbe portato alla stabilità (*homeostasis*). Venivano infatti presentati i principi generali per progettare un “amplificatore di selezione” (*selection amplifier*), che avrebbe velocizzato il processo in modo da raggiungere soluzioni non contemplate dal progettatore, ottenendo quindi un’ “amplificazione” dell'intelligenza. Secondo la concezione di Ashby, originalità e produttività non erano caratteristiche essenziali dell'intelligenza, ma delle informazioni casuali fornite a un selezionatore, ovvero colui che operava la scelta: era la capacità di operare una selezione appropriata che riconduceva la produttività e l'originalità a un processo

¹¹¹ Ashby 1956: 216-217.

altamente disciplinato che avrebbe portato verso il risultato voluto.¹¹²

Ashby scrisse a McCarthy:

I am afraid I cannot consent to having it (il suo articolo) “refereed”, as it is on a subject to which I have devoted a great deal of work, and on which I can now speak with some authority. Frankly, I know of no referee to whose opinion I would defer (though I would, of course, be prepared to answer specific objections).¹¹³

Dimostrando ben più modestia, MacKay accettò invece i *referee report* per il suo articolo. Alla fine, McCarthy e Shannon accettarono gli articoli di Ashby e MacKay senza operare nessun *refereeing*. L’esperienza probabilmente non giovò all’opinione che McCarthy aveva di Ashby, veterano della cibernetica inglese.

Shannon e McCarthy divisero gli articoli in tre sezioni: *Finite Automata* (automi con un numero limitato di possibili stati interni, come le reti neurali); *Turing Machines* (automi con un illimitato numero di possibili stati interni); e *Synthesis of Automata* (automi che simulavano il funzionamento di un organismo vivente). Queste categorie riprendevano in parte la divisione concettuale già operata da Shannon nel suo articolo

¹¹² Asaro 2008: 170-173.

¹¹³ 16 giugno 1954.

Computers and Automata del 1953.¹¹⁴ Più in generale, tale distinzione rappresentava la sua visione generale sull'argomento, espressa nel 1958:

The theory of automata is a relatively recent and by no means sharply defined area of research. It is an interdisciplinary science bordered mathematically by symbolic logic and Turing machine theory, bordered engineering-wise by the theory and the use, particularly for general non-numerical work, of large scale computing machines, and bordered biologically by neurophysiology, the theory of nerve-nets and the like.¹¹⁵

L'unica categoria mancante negli *Automata Studies* era proprio la manipolazione simbolica tramite calcolatori, sebbene a essa si accenni, solo brevemente, nella prefazione del libro:

Developments in symbolic logic, recursive function theory and Turing machine theory have led to deeper understanding of the nature of what is computable by machines.¹¹⁶

Di contro, la cibernetica era invece molto ben rappresentata: quasi la metà del volume (6 articoli su 13) erano sulle teorie e sull'uso delle reti neurali, e praticamente tutta la terza parte era curata dai cibernetici britannici (Ashby, MacKay e Uttley).

Forse proprio per questi motivi, a lavoro concluso, la reazione di McCarthy fu di estrema delusione. A metà 1954 scrisse a Shannon, con una punta di

¹¹⁴ Il quale trattava essenzialmente di reti neurali, macchine di Turing e macchine capaci di replicare attività tipicamente umane.

¹¹⁵ Shannon 1958: 831.

¹¹⁶ McCarthy, Shannon 1956: 728.

amarezza, che la raccolta nella sua interezza non costituiva un grosso passo in avanti rispetto agli studi del tempo, ma era certamente rappresentativa del paradigma dominante. Le aspettative di McCarthy erano alte, e forse in parte irrealizzabili. Egli pensava che gli articoli ad opera del Ratio Club “contribute heuristic ideas of some value but none lead directly to a solution of the problem of thinking automata”, mentre quelli di Kleene e James Culbertson sulle reti neurali “solve problems of mainly formal interest”. McCarthy non risparmiava nemmeno l’autocritica. Il suo articolo per la raccolta, *The Inversion of Functions Defined by Turing Machines*, introduceva il concetto di problema ben definito:

We call a problem well-defined if there is a test which can be applied to a proposed solution. In case the proposed solution is a solution, the test must confirm this in a finite number of steps.¹¹⁷

Anche McCarthy si rifaceva alla definizione di computabilità usata da Turing: se la soluzione di un problema intellettuale ben definito può essere trovata tramite un numero finito di stadi, allora questo problema poteva essere considerato come la funzione inversa di quello definito da una macchina di Turing. Ma invertire funzioni computabili non era, in linea di principio, il modo migliore per raggiungere i suoi scopi, e McCarthy parve accorgersene quasi subito: riguardo il suo articolo scrisse infatti che esso introduceva “concepts which may prove useful but does not get far off the

¹¹⁷ McCarthy 1956.

ground”. Il problema era di carattere sensibile: piccoli cambiamenti nella struttura della macchina comportavano enormi modificazioni nel comportamento.

Il contributo di Shannon agli *Automata Studies* era costituito dall’articolo *A Universal Turing Machine with Two Internal States*. In esso, Shannon dimostrava come fosse possibile progettare una macchina di Turing universale (ovvero capace di comportarsi come qualunque macchina di Turing specifica, quando provvista della descrizione di quest’ultima) dotata solamente di un nastro infinito e due stati interni per operare. Per fare ciò, era necessario aumentare il numero di simboli impressi sul nastro, visto che lo stato della macchina originale (ovvero quella il cui comportamento la macchina di Turing universale avrebbe dovuto emulare) doveva essere depositato e aggiornato sul nastro dell’altra. In questo modo, concludeva Shannon, sarebbe stato possibile costruire una macchina di Turing universale dotata di due soli stadi interni. Egli dimostrava anche come fosse impossibile costruire una macchina analoga dotata di un solo stato interno. In questo caso il giudizio di McCarthy fu senza dubbio più positivo: “A short step in the direction of the theory of computing rate (delle macchine di Turing)”. Ma fu l’articolo di von Neumann, *Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components*, a ottenere la valutazione migliore. Von Neumann partiva dal presupposto che l’organismo fosse capace di auto-diagnosticare i propri malfunzionamenti e

di limitarne gli effetti: era per questo che i sistemi naturali non necessitavano di interventi esterni. Il cervello degli animali e dell'uomo, pur essendo costituito da strutture fragili e poco affidabili come i neuroni e le sinapsi, continuava a funzionare anche quando si verificano danni più o meno gravi. Le macchine artificiali, invece, erano progettate in modo che ogni errore si amplificasse, rendendo necessaria l'identificazione rapida del problema per la riparazione o la sostituzione del componente difettoso.

Assumendo quindi che sia il cervello umano sia le macchine erano entrambi costituiti da elementi poco affidabili, per rendere le prestazioni degli automi simili a quelli degli umani era necessario costruire codici per la trasmissione d'informazioni in modo che questa fosse affidabile anche quando la trasmissione dei segnali e dei simboli era soggetta a errori.

McCarthy scrisse a Shannon: "An important result, perhaps slightly off the main track".¹¹⁸

Pare invece che Minsky abbia ritardato parecchio nel presentare il suo articolo, ed è probabilmente per questo che non si trovano commenti a riguardo.¹¹⁹

Non sorprende, invece, che l'opinione di Shannon sulla raccolta di saggi fosse molto più positiva. Oltre al suo interesse generale per la cibernetica (per non parlare degli ottimi rapporti che aveva con i suoi principali esponenti), Shannon in passato aveva speso parole di elogio per l'omeostato

¹¹⁸ 11 agosto 1954.

¹¹⁹ 18 gennaio 1955.

di Ashby, definendolo “a basis for learning machines and brain modelling”.¹²⁰

Le tensioni in fase di editing riemersero in una lettera che McCarthy scrisse a Shannon nel tardo 1954. McCarthy aveva scritto una prima bozza della premessa del volume:

It consists mainly of a point of view on how the various lines of investigation represented by the included papers may contribute to the eventual design of intelligent machines¹²¹.

La frase è sottolineata da una linea ondulata, forse da Shannon stesso.¹²²

McCarthy aveva anche altri dubbi:

1. Do you agree that the book should be regarded as a step in the direction of the design of intelligent automata? Or is this too presumptuous? Might it surprise the contributors?
2. What about the title “Towards Intelligent Automata”? Is it too bombastic? Others?

Probabilmente questi dubbi derivavano dal fatto che Shannon trovasse l'interpretazione del volume da parte di McCarthy troppo presuntuosa e il titolo da lui proposto troppo altisonante. La prefazione pubblicata nel volume de-enfatizzava lo scopo di progettare “automi intelligenti” e si concentrava, invece, sul funzionamento del cervello e sulla possibilità di

¹²⁰ Shannon 1953: 709.

¹²¹ 24 novembre 1954.

¹²² Kline 2001: 8.

simularlo tramite una macchina, definendo entrambi i punti come “analytic and synthetic problems”. Piuttosto che di intelligenza, McCarthy e Shannon sembravano aggirare la questione parlando di “pensiero” (*thinking*), riprendendo la definizione che ne aveva dato Turing. Quest’ultima aveva il vantaggio di essere “operational or, in the psychologists' term, behavioristic”, ma aveva anche dei forti limiti concettuali:

Such a machine, in a sense, for any given input situation (including past history) merely looks up in a "dictionary" the appropriate response. With a suitable dictionary such a machine would surely satisfy Turing's definition but does not reflect our usual intuitive concept of thinking. This suggests that a more fundamental definition must involve something relating to the manner in which the machine arrives at its responses--something which corresponds to differentiating between a person who solves a problem by thinking it out and one who has previously memorized the answer.¹²³

Il titolo rimase quello ben più posato di *Automata Studies*.

A una prima analisi, può quasi sembrare che l’influenza di Shannon abbia assunto più che altro la funzione di “blocco” nel raggiungimento del risultato che McCarthy si era prefissato. In realtà, Shannon simpatizzava molto per l’obiettivo di McCarthy, tanto da scrivere, in una lettera indirizzata a Irene Angus (suo professore di scienze ai tempi del liceo):

“My fondest dream is to someday build a machine that really thinks, learns, communicates with humans and manipulates its environment in a fairly

¹²³ McCarthy, Shannon 1956: 729.

sophisticated way”.¹²⁴ Shannon approvò inoltre la pubblicazione degli articoli scritti dal Radio Group sui principi necessari per la progettazione di automi intelligenti.

Nonostante tutto, McCarthy non era contrario a menzionare la cibernetica nella premessa del libro. In una lettera a Shannon, essa veniva descritta come una delle aree di ricerca cui la prefazione poteva accennare (*brief bows*), insieme alla logica matematica, la statistica, la teoria della comunicazione, la neurofisiologia, e l'ingegneria telefonica.¹²⁵ La prefazione pubblicata, però, firmata da McCarthy e Shannon, menzionava tutte queste discipline e i loro principali esponenti, ad eccezione dei cibernetici. Ancora oggi non si sa chi fu a cancellare questi riferimenti. In conclusione, il primo tentativo di fondazione dell'AI non si rivelò molto produttivo. La maggior parte degli scritti ricevuti per il libro si concentrava infatti sulla teoria degli automi nel senso più stretto del termine, cioè sui principi matematici che erano alla base del funzionamento dei sistemi elettromeccanici, e non sulla relazione tra linguaggio e intelligenza, o sulla capacità delle macchine di giocare, sulla programmazione, e in definitiva su quegli argomenti che esercitavano così tanto fascino su McCarthy.

L'esperienza, comunque, gli impartì un'importante lezione.

Il nome *Automata Studies* non era il più idoneo per rappresentare il tipo di indagini che McCarthy aveva in mente: molti degli articoli ricevuti non

¹²⁴ 8 agosto 1952.

¹²⁵ 24 novembre 1954.

avevano nulla a che fare con la riproduzione dell'intelligenza umana, sebbene fosse questo il tema portante del libro (cosa che, peraltro, veniva rimarcata nella prefazione del libro, per quanto in maniera non eccessivamente diretta). Così, nella sua proposta e poi durante la stessa Conferenza di Dartmouth, egli argomentò con forza a favore del termine "artificial intelligence":

As for myself, one of the reasons for inventing the term "artificial intelligence" was to escape association with "cybernetics". Its concentration on analog feedback seemed misguided, and I wished to avoid having either to accept Norbert (not Robert) Wiener as a guru or having to argue with him.¹²⁶

Come suggerisce l'aggettivo "artificiale", lo scopo primario della nascente AI (almeno secondo il punto di vista di McCarthy) non era quello di studiare e comprendere il comportamento umano. In questo senso, la cibernetica non era un approccio sbagliato in sé: semplicemente, i suoi metodi e scopi erano diversi e improduttivi per gli obiettivi di McCarthy. Inizialmente, l'intenzione era quella di prendere in considerazione il comportamento umano solo nella misura in cui questo poteva fornire un modello efficace nella progettazione delle macchine, affinché svolgessero determinati compiti:

For this reason, whatever revolution there may have been around the time of the Dartmouth Project was to get away from studying human behavior and to consider

¹²⁶ McCarthy 1989.

the computer as a tool for solving certain classes of problems. Thus AI was created as a branch of computer science and not as a branch of psychology. Newell, Simon and many of their students work both in AI as computer science and AI as psychology.¹²⁷

Il punto non era costruire macchine che replicassero, strutturalmente e funzionalmente, il cervello umano e i comportamenti intelligenti che scaturiscono dalle sue funzioni, quanto piuttosto progettare macchine in grado di esibire un comportamento che in certi contesti avrebbe potuto definirsi intelligente, prendendo i processi cognitivi come esempi di procedure efficaci.

Nacquero anche alcune controversie sul nome che McCarthy aveva scelto per la nuova disciplina. Per Samuel, ad esempio, il termine “artificiale” faceva pensare a qualcosa di contraffatto, irreali. Il termine non piacque né a Newell né a Simon che, negli anni seguenti, chiamarono il loro lavoro “elaborazione di informazioni complesse” (*complex information processing*).¹²⁸ Questo almeno fino al 2005, quando Newell sembrò averlo ormai accettato definitivamente:

So cherish the name artificial intelligence. It is a good name. Like all names of scientific fields, it will grow to become exactly what its field comes to mean.¹²⁹

¹²⁷ Ibid.

¹²⁸ McCorduck 1987: 115.

¹²⁹ Newell 2005: 29.

McCarthy trascorse l'estate del 1955 lavorando all'IBM; fu in questo periodo che iniziò a concentrare la propria attenzione sui calcolatori digitali. Il lavoro manuale su di essi lo convinse che questi dispositivi rappresentavano il mezzo migliore per realizzare il suo progetto, al contrario della macchina di Turing e la teoria degli automi, che si limitavano a sfruttare l'intelligenza umana in maniera astratta e poco costruttiva.¹³⁰

2.5 Il finanziamento per il seminario

Sulla base di queste conclusioni, McCarthy decise di contattare la Rockefeller Foundation, al fine di ottenere i fondi necessari per l'organizzazione di un seminario avente come oggetto l'AI. Arrivato al Dartmouth College nel febbraio 1955 come assistente di matematica, McCarthy inviò a Shannon la versione definitiva del suo articolo per gli *Automata Studies*, aggiungendo di essere ora disponibile per eventuali revisioni.¹³¹ Nell'aprile di quell'anno, McCarthy ebbe un colloquio con Warren Weaver, il direttore della divisione di scienze naturali della Rockefeller Foundation, riguardo la possibilità di finanziare un seminario estivo di sei settimane in cui 10 partecipanti avrebbero lavorato insieme sulla teoria del cervello e degli automi. Weaver trascrisse sulla sua agenda lavorativa, che circolò all'interno della fondazione, che McCarthy era sicuro che Shannon e Rochester dell'IBM avrebbero partecipato al

¹³⁰ Crevier 1993: 38.

¹³¹ 18 febbraio 1955.

seminario e che i rispettivi datori di lavoro erano disposti a pagare i loro stipendi durante la permanenza a Dartmouth. Tra i possibili partecipanti, McCarthy aveva fatto anche il nome di Minsky, e avrebbe voluto che personalità di spicco come von Neumann e George Polya partecipassero al seminario per almeno qualche giorno. Weaver rispose che ci avrebbe pensato e che si sarebbe fatto sentire presto.¹³²

In realtà, Weaver provava sentimenti contrastanti nei confronti del progetto. Poco dopo il primo colloquio, disse a McCarthy che sarebbe stato Robert Morison, diretto della divisione di ricerca biologica e medica della fondazione, a occuparsi della proposta, probabilmente per via del suo interesse verso il *brain modelling*.¹³³ Non potendo partecipare al pranzo programmato per metà giugno con Morrison, McCarthy e Shannon (che nel frattempo era stato “arruolato” da McCarthy per promuovere il progetto), Weaver espresse a Morison la sua opinione sulla faccenda.

Studio di matematica applicata con un certo interesse verso la meccanica, molte delle ricerche sui primi calcolatori furono finanziate dalla Rockefeller Foundation proprio sotto la supervisione di Weaver. Tra questi progetti vi era anche l’analizzatore differenziale di Bush. Weaver inizialmente era restio a finanziare le ricerche di Bush, e si convince solo dopo molti anni, quando quest’ultimo avanzò la possibilità che la sua macchina potesse

¹³² Tratto dall’agenda personale di Weaver, 4 Aprile 1955 . Il documento, insieme agli altri che riguardano direttamente la Rockefeller Foundation, è consultabile presso il sito: <http://rockefeller100.org>.

¹³³ 7 giugno 1955.

aprire nuovi campi di ricerca scientifica, come quello della “biologia matematica”.¹³⁴ Visti i precedenti, è probabile che la fondazione fosse in quel periodo più interessata a progetti “utili” e poco rischiosi, già avviati e che non necessitassero di nessuna elaborazione teorica. Weaver sospettava che il seminario estivo fosse più che altro un “progetto personale”, nel senso che McCarthy e gli altri partecipanti si sarebbero divertiti a passare un po’ di tempo insieme, discutendo di vari aspetti della teoria dell’informazione: quasi un tranquillo meeting intellettuale informale, senza nessun sviluppo concreto e a spese di qualcun altro. Il progetto era “personale” anche perché non vi era nessun gruppo o organizzazione ad appoggiarlo (“I would, for example, think of it in rather different terms if the idea came to us from the entire information group at MIT”).

Viste le premesse, non vi erano le basi per concedere un finanziamento. Di contro, Weaver nutriva la massima stima per Shannon. I due avevano già collaborato in passato. Nel 1948, Shannon pubblicò l’articolo *A Mathematical Theory of Communication*, in cui dimostrava che tutte le fonti d’informazioni - telegrafi, persone che parlano, telecamere e così via - avevano un “tasso di fonte” (*source rate*) a essi associato che poteva essere misurato in bit al secondo. I canali di comunicazione erano dotati di una “capacità” misurabile nelle stesse unità: le informazioni potevano quindi essere trasmesse sul canale se e solo se il tasso di fonte non superava la capacità del canale. Ristampato l’anno successivo, il titolo dell’articolo

¹³⁴ Mindell 2002: 170.

venne cambiato in *The Mathematical Theory of Communication*, probabilmente per sottolineare che le conclusioni di Shannon avrebbero dovuto essere prese come punto di riferimento per tutti gli studi futuri sull'argomento. La nuova edizione, oltre l'articolo di Shannon (rimasto sostanzialmente invariato), presentava anche dei capitoli aggiuntivi scritti da Weaver, il quale figura come co-autore. Weaver non solo conosceva Shannon, ma contribuì anche a rendere popolari le sue ricerche sulla teoria dell'informazione.

Dimostrando un'onestà d'altri tempi, Weaver temeva il conflitto d'interessi, tanto da scrivere a Morison:

I think it is on the whole rather better that I am not going to be at the luncheon. My personal connection with the subject, especially since Claude Shannon is now going to be present, tends to remove the discussion from the objective atmosphere in which we ought to approach it.¹³⁵

Ma pare che fu proprio la presenza di Shannon a rendere Morison più possibilista nei confronti del progetto. Nella sua agenda, egli trascrisse un breve resoconto dell'incontro. Morison esordì scusandosi per l'apparente scetticismo della Rockefeller Foundation nei confronti del progetto: era difficile valutare qualcosa come un meeting di ricerca sulla base delle sole intenzioni, e ciò rendeva necessario analizzare ogni richiesta più che scrupolosamente. Egli pose anche una domanda abbastanza atipica ma per

¹³⁵ Whisler 1970: 11.

nulla scontata: in quel determinato periodo storico, vi era davvero bisogno di un confronto tra scienziati per far nascere nuove idee? O sarebbe stato meglio incentivare gli sforzi del singolo che, lavorando da solo, si concentrava unicamente sui propri studi? L'esempio usato da Morison era quello degli studi neurologici in connessione con il metodo matematico, ancora fermi alla loro fase pre-newtoniana: forse in quel momento la scienza aveva più bisogno di intuizioni brillanti che di discussioni di gruppo.

McCarthy e Shannon erano in parte d'accordo, ma specificavano anche che la maggior parte delle persone che avrebbero voluto invitare alla conferenza disponevano di pochissimo tempo libero per l'elaborazione teorica delle loro idee. Uno degli obiettivi del seminario sarebbe stato proprio quello di dare a ognuno dei partecipanti un periodo abbastanza lungo per dedicarsi a tempo pieno allo sviluppo teorico delle ricerche, con in più lo stimolo costituito dagli incontri occasionali e dallo scambio continuo d'idee.

Morison era ancora titubante: l'idea aveva senza dubbio la sua logica, ma forse lo era un po' meno che la Rockefeller Foundation vi contribuisse.

I partecipanti proposti per il seminario erano McCarthy, Shannon, Rochester, Minsky, von Neumann (per un breve periodo, se gli impegni glielo avessero permesso) e Donald Hebb.

Morison aveva già finanziato, insieme alla divisione di Weaver, le ricerche matematiche e neuropsicologiche di Wiener e Rosenblueth a partire dalla

metà degli anni Quaranta, ed è probabile che volesse includere anche qualche ricercatore proveniente dal gruppo di ricerca di Wiener al MIT: si può giustificare così la presenza di Selfridge nella lista dei possibili partecipanti. Morison scrisse che Selfridge, in quel periodo *group leader* dei Lincoln Lab al MIT, “apparently strike M. (McCarthy) and S. (Shannon) as the soundest and most amenable to group activity of the various younger men who have been associated with Norbert Wiener”.

Morison stava probabilmente alludendo alle difficoltà che Wiener e Rosenblueth ebbero con il brillante ma abbastanza imprevedibile Pitts durante le ricerche finanziate dalla Rockefeller Foundation.¹³⁶ Ma è anche probabile che sia stato Minsky a raccomandare Selfridge, dal momento che egli face parte, per qualche del tempo, del suo gruppo di ricerca ai Lincoln Lab. Fu in quel periodo che scoprirono di condividere gli stessi interessi in materia d’intelligenza meccanica, tanto da dirigere insieme un meeting sull’apprendimento artificiale durante la *Western Joint Computer Conference* nel 1955, un anno prima la Conferenza di Dartmouth.¹³⁷

Durante il pranzo venne considerata la partecipazione di altri due cibernetici. Nonostante le difficoltà avute con Ashby durante la lavorazione degli *Automata studies*, McCarthy e Shannon lo definirono “the most attractive foreign name”. Shannon, invece, nutriva qualche dubbio su

¹³⁶ Ancora oggi le cause che portarono alla brusca interruzione della collaborazione tra Wiener e Pitts sono oggetto di speculazione. Riguardo Selfridge e gli altri giovani ricercatori appartenuti al gruppo di Wiener, si veda anche Conway, Siegelman 2005: 140-142 e 201-204.

¹³⁷ Husbands 2008b: 399.

MacKay. Quest'ultimo non era certo un nome nuovo per Morison, dal momento che arrivò in America nel 1951 grazie a una borsa di studio della Rockefeller Foundation, la quale avrebbe, da lì in poi, finanziato le sue ricerche sulla teoria dell'informazione e sugli automi analogici.¹³⁸ In un articolo sempre del 1951, *Mindlike Behaviour in Artefacts*, Mackay proponeva un meccanismo statistico che avrebbe presentato molte delle qualità dei processi cognitivi umani, compresi i pregiudizi, le preferenze, l'originalità e l'apprendimento. Scelse di chiamare tale comportamento "simile a quello della mente" (*mindlike*), ma la particolarità di quest'articolo stava anche nel dedicare una sezione apposita ai problemi prettamente filosofici da esso sollevati, come quello del rapporto mente-corpo. Per MacKay molte delle osservazioni fatte in merito si basavano su falsi dualismi nati dalla domanda "how can matter produce mind?", in cui era il presupposto logico a essere sbagliato. Non era possibile astrarre il sistema dal suo ambiente (dal momento che molte caratteristiche del comportamento nascevano dalla loro interazione), e non era possibile ricondurre le caratteristiche del pensiero a "nient'altro che" ("nothing but"¹³⁹) astrazioni delle descrizioni fisiche, dal momento che appartengono a due gruppi logici differenti (ed è per questo motivo che sembravano apparentemente inconciliabili). Era possibile, invece, basarsi su un dualismo fondato su due cornici concettuali di riferimento, diverse ma

¹³⁸ McCorduck 1987: 96.

¹³⁹ Le virgolette sono presenti nel testo originale.

complementari. Come le descrizioni duali della luce intesa come onde e come particelle, ogni descrizione era valida nel proprio ambito, senza il bisogno di spiegarne una come se fosse “nient’altro che” una conseguenza dell’altra.

Prendendo le mosse da questa distinzione fondamentale, MacKay dimostrava come concetti quali coscienza, auto-coscienza, libero arbitrio e Io fossero applicabili anche al suo “artefatto”, fornendo per essi delle nuove definizioni abbastanza originali.

MacKay si premurava anche di ribattere a ipotetiche obiezioni di carattere teologico:

No reputable theologian expects to find physical laws disobeyed in the human brain; and it is difficult to see how elucidation of the particular physical processes which happen to be used can in any way be relevant to the claims (whether admitted or not) which are made by Christianity on the personality whose thought is mediated by those processes.¹⁴⁰

Fede e scienze erano complementari, non antagonisti, esattamente come lo erano la mente e il corpo. Eppure, era proprio questo interesse di Mackay per la teologia a non convincere Shannon. Non è chiaro se si trattasse di una semplice divergenza di vedute: se, cioè, Shannon non condividesse le convinzioni di MacKay, e che quindi mettesse in discussione il suo possibile contributo alla conferenza. Restava comunque il fatto che la

¹⁴⁰ MacKay 1951: 120.

religiosità di MacKay avrebbe potuto far nascere degli attriti tra i partecipanti, in quanto scienziati generalmente poco propensi ad accettare una prospettiva del genere.

Morison chiese anche quanto ci sarebbe voluto per passare dalla fase di pura speculazione sulle capacità dei calcolatori all'analogia diretta con le funzioni cerebrali. McCarthy e Shannon risposero che si trattava di una questione delicata, e che sarebbe stato meglio "to go rather slowly at this".

Evidentemente la risposta non piacque, tanto che Morison suggerì di includere anche due psicologi "for the purpose of keeping the group from speculating too wildly on how the brain might work". La scelta ricadde su Hans Lukas Teubert, della New York University, che aveva co-diretto gli atti delle Conferenze Macy sulla cibernetica, e Karl Pribram. Morison pensava che entrambi fossero "enough breadth of education and curiosity to have informed themselves, at last fairly well, about cybernetics and mathematical speculations"¹⁴¹ – uno dei pochi riferimenti diretti alla cibernetica durante l'organizzazione della Conferenza di Dartmouth.

In definitiva, il giudizio di Morison sul progetto era tutt'altro che positivo. Egli dubitava che potesse uscirne qualcosa di buono a costi ragionevoli, ed era ancora più scettico sulla possibilità che la maggior parte degli invitati vi presenziassero davvero. Ironicamente, l'unica partecipazione certa, ovvero quella di MacKay, era anche la più sgradita. Il vero discrimine sarebbe stato la proposta formale che McCarthy e Shannon avrebbero presentato, nel caso

¹⁴¹ 17 giugno 1955.

in cui essa fosse stata in grado di suscitare o meno interesse della
fondazione.

3. La Conferenza di Dartmouth

Con le premesse illustrate nel capitolo precedente, McCarthy riunì un piccolo gruppo di scienziati per la stesura della Proposta di Dartmouth, comprendente sé stesso, Shannon, Rochester e un'altra figura che finora è comparsa a più riprese e che si rivelerà fondamentale per il futuro l'AI: Marvin Minsky.

3.1 Marvin Minsky

Nello stesso periodo in cui Hebb sviluppava la teoria dell'apprendimento cellulare, un insolito studente di Harvard giungeva in maniera indiretta a conclusioni simili. Minsky era un allievo della Bronx High School of Science poi passato ad Harvard dove, come disse ironicamente il suo amico e collega di università Jeremy Bernstein, non era del tutto chiaro quale fosse la sua specializzazione o se ne avesse scelta una. Formalmente era uno studente di fisica, ma durante il primo anno di università Minsky convinse lo zoologo John Welsh a farlo lavorare per conto proprio presso il laboratorio di biologia. Gli venne data una grande stanza ricca di attrezzatura, dove divenne un esperto della neurofisiologia degli astaci (i gamberi di fiume): sezionandoli e collegandone i nervi a una sorgente elettrica, Minsky ne faceva muovere le chele in modo che afferrassero una matita e la mettessero dentro un vaso.¹⁴² Nei momenti in cui non studiava

¹⁴² Bernstein 1990: 17.

fisica e non lavorava con i suoi gamberi di fiume, Minsky cominciò a frequentare anche il laboratorio di psicologia. Qui ebbe modo di farsi un'idea abbastanza precisa sulle diverse correnti di pensiero che caratterizzavano la psicologia di fine anni Quaranta.

Un'estremità del laboratorio era occupata dallo psicologo comportamentista Burrhus F. Skinner. La nascita del comportamentismo viene solitamente fatta coincidere con la pubblicazione di *Animal Intelligence* ad opera di Edward Lee Thorndike. Esso consisteva in una brutale trasposizione degli esperimenti Pavloviani condotti sugli animali alla psicologia umana. Uno di essi consisteva, ad esempio, nel puntare una luce verso un gatto nel momento in cui gli veniva portato del cibo. Dopo qualche giorno, il gatto avrebbe iniziato a salivare in risposta alla luce anche in assenza di cibo. Pavlov chiamava la luce "stimolo" (*stimulus*) e la salivazione "risposta" (*response*). Per i comportamentisti, tutte le azioni, i pensieri, i desideri, e in genere qualunque caratteristica dell'essere umano, erano semplici riflessi attivati da una forma più "alta" di stimoli. L'unica differenza tra animali ed esseri umani consisteva nel fatto che questi ultimi reagivano a degli stimoli più complessi, chiamati "situazioni" (*situations*). Visto che la mente aveva la sola funzione di associare situazione e risposte, non aveva senso esaminarla usando, ad esempio, l'introspezione. Simon, pioniere della psicologia cognitiva, avrebbe detto che in quel periodo era impensabile usare parole come "mente" negli articoli psicologici, altrimenti si correva il

rischio di “get your mouth washed out with soap”, come quando un bambino dice una parolaccia.¹⁴³ Per certi estremisti, come lo era Skinner, la mente non esisteva affatto. Negli anni Cinquanta si venne a creare quella situazione paradossale in cui, mentre gli psicologi quasi rigettavano le fondamenta storiche della loro disciplina, gli ingegneri e i matematici, di contro, costruivano macchine che giocavano a scacchi, che provavano teoremi matematici, che contenevano al proprio interno una memoria, discutendo delle “menti” di questi sistemi in maniera libera e senza nessun condizionamento. Minsky era scettico nei riguardi del comportamentismo; di contro, provava una grande stima nei confronti di Skinner, tanto da aiutarlo in alcuni dei suoi esperimenti. E’ possibile che le idee di Skinner sull’apprendimento per rinforzo abbiano successivamente convinto Minsky a costruire una rete neurale.¹⁴⁴

Ma i comportamentisti non erano i soli a occupare il laboratorio di psicologia:

A occidente vi erano i comportamentisti, che tentavano di campire il comportamento senza una teoria; a oriente c’erano gli psicofisiologi, che tentavano di capire un minuscolo pezzo di sistema nervoso senza avere alcuna idea del resto. [...] Gli specializzandi [...] si radunavano nel mezzo dello scantinato e discutevano dell’una e dell’altra dottrina.¹⁴⁵

¹⁴³ Holden 1986.

¹⁴⁴ Crevier 1993: 33.

¹⁴⁵ Bernstein 1990: 41.

Tra questi specializzandi vi era anche George Miller, che in quel periodo cercava di elaborare delle teorie sul funzionamento della mente tramite il formalismo matematico. Qualche anno dopo, nel 1956, Miller sarebbe diventato famoso grazie a una pubblicazione incentrata sulla memoria a breve termine. Intitolato *The Magical Number Seven*, l'articolo gettava una luce critica sui processi cognitivi umani. L'essere umano non era in grado, secondo Miller, di tenere a mente più di sette informazioni di diversa provenienza. Nel porre questa limitazione, Miller enfatizzava il ruolo attivo della mente come elaboratore d'informazioni, distaccandosi nettamente dal modello comportamentista di pura associazione meccanica passiva. Fu con Miller che Minsky decise di studiare l'apprendimento:

I found this thing called the *Bulletin of Mathematical Biophysics*. It had works of Warren McCulloch and the great pioneers of the 1940's... As soon as I saw that I started to think "How could I make a learning machine?"¹⁴⁶

Laureatosi ad Harvard con una tesi sulla topologia, Minsky iniziò il suo dottorato di ricerca a Princeton, dove ebbe modo di conoscere McCarthy. Già in quel periodo i due ebbero, per usare le parole di McCarthy, delle conversazioni sull'intelligenza meccanica "non molto produttive, che erano servite solo a stabilire che eravamo di fatto allineati in questo campo e che concordavamo su un certo numero di cose – non credo che Minsky e io abbiamo modificato molto l'uno le idee dell'altro sui punti su cui non

¹⁴⁶ Crevier 1993: 34.

eravamo d'accordo, ma, sui punti su cui eravamo d'accordo, ci rafforzammo reciprocamente".¹⁴⁷ E' facile intuire quali fossero i punti di disaccordo: le ricerche di Minsky durante i primi anni a Princeton erano ancora pesantemente influenzate dal lavoro di McCulloch e Pitts, e probabilmente anche dal comportamentismo di Skinner. Anche anni dopo, Minsky non avrebbe mai negato questo suo iniziale "innamoramento" per le reti neurali.¹⁴⁸ Durante l'estate del 1951, insieme a Dean Edmonds (suo collega e studioso di elettronica), Minsky costruì la sua prima rete neurale usando trecento valvole termoioniche e il pilota automatico di un B-24. Questa consisteva in una rete di quaranta neuroni artificiali che simulavano il cervello di un topo intento a trovare l'uscita da un labirinto. A ogni neurone corrispondeva una parte del labirinto e, quando attivato, esso avrebbe rappresentato la "consapevolezza" del topo di trovarsi in quel determinato punto. I neuroni connessi con quello attivato rappresentavano invece le scelte che il topo avrebbe potuto fare come, per esempio, se spostarsi a destra o sinistra. Quale di questi neuroni si sarebbe attivato dipendeva dalla forza delle loro connessioni con il neurone già attivo, ed era compito del pilota automatico regolarle. Al progetto collaborò anche Miller, il quale aveva ottenuto un finanziamento di 2000 dollari dall'Office of Naval Research.

La meccanica della rete neurale di Minsky riprendeva in parte quella degli

¹⁴⁷ McCodruck 1987: 112.

¹⁴⁸ Norberg 1989: 10.

stimoli-ricompensa tipica del comportamentismo. Minsky provò a parlare della sua macchina a Skinner, il quale, tuttavia, non si dimostrò molto interessato. Presto gli apparve evidente che le tecniche di apprendimento skinneriane non l'avrebbero portato da nessuna parte: esse non offrivano alla macchina nessun modo di ragionare davvero su cosa stesse effettivamente facendo, e quindi di prendere le decisioni sulla base delle conclusioni ottenute.

Minsky conseguì il suo dottorato nel 1954, convinto che una rete neurale, per rispecchiare fedelmente il comportamento del cervello umano, avrebbe richiesto migliaia e forse milioni di neuroni artificiali. Costruire una rete del genere era un'impresa fuori portata, e ciò lo convinse a cercare altri metodi per creare macchine pensanti.

Minsky tornò successivamente ad Harvard in qualità di *junior fellow*, su richiesta di Andrew Gleason e con il benestare di von Neumann, Wiener e Shannon. Qui ebbe modo di continuare per tre anni le proprie ricerche in completa libertà, dedicandosi temporaneamente all'ottica e inventando, nel 1955, il primo microscopio confocale.¹⁴⁹

E' durante questo periodo che avvenne il primo incontro con Ray Solomonoff:

I met a young man named Ray Solomonoff who was working on an abstract theory of deductive inference. He had worked on a learning machine... that was pretty formal. I was so impressed I decided this was much more productive than the neural net system, in which you built a piece of hardware and hoped it would

¹⁴⁹ Minsky 1988: 128.

do the right thing. With this new approach, you tried to make theories of what kind of inference you wanted to make, and then asked, “How would I make a machine do exactly that?”. It was a different line of thought¹⁵⁰.

In seguito a queste conversazioni, Minsky iniziò progressivamente ad afferrare la differenza fondamentale tra sapere com'è costituito il cervello e capire quello che effettivamente fa. I calcolatori digitali stavano iniziando a mostrare un modo nuovo per percorrere quest'ultima via. Ora era possibile per gli scienziati programmare i calcolatori affinché rispecchiassero le loro teorie sul funzionamento della mente, e fare in modo che essi si comportassero esattamente in quel modo.

Il 1955 fu sicuramente un anno ricco di avvenimenti per Minsky, visto che risale a quell'anno anche la stesura della Proposta di Dartmouth.

3.2 La Proposta di Dartmouth

McCarthy, Shannon, Minsky e Rochester finirono la stesura delle rispettive sezioni di ricerca per la proposta nel tardo agosto del 1955, ed essa venne inviata alla Rockefeller Foundation ai primi di settembre di quello stesso anno¹⁵¹. Citando Paul Edwards: “In the proposal for work at conference, the still-nascent split between the computer-brain and computer-mind

¹⁵⁰ Crevier 1993: 37. La teoria di Solomonoff, riscoperta indipendente nel corso degli anni da altri ricercatori, è oggi chiamata “teoria della probabilità algoritmica”.

¹⁵¹ McCarthy conobbe Rochester durante l'estate del 1955: “In the summer of 1955 I had a summer job at IBM (I was at Dartmouth at the time). I was hired by Nathaniel Rochester. Rochester hired me because IBM had given, or was about to [and] had committed itself to [give] this 704, do so to MIT and the New England universities on the basis that it would be located at MIT [...] So I worked in his information research department and turned Rochester on to AI”. Cfr. Aspray 1989: 8.

metaphors already appears clearly”.¹⁵² Tuttavia, essa includeva entrambi gli approcci, considerandoli come metodi alternativi e complementari nell’impresa comune di voler utilizzare i calcolatori come strumenti di ricerca. L’introduzione, probabilmente scritta da McCarthy, descriveva gli “aspetti problematici dell’intelligenza artificiale” (*artificial intelligence problem*), dividendoli in sette categorie: calcolatori automatici, come programmare un calcolatore per manipolare un linguaggio, reti neurali, teoria delle dimensioni di un calcolo, auto-miglioramento, astrazioni, casualità e creatività. La sezione dedicata al rapporto linguaggio-intelligenza, una delle aree di ricerca cui McCarthy era maggiormente interessato, era chiaramente riconducibile a quella che successivamente sarebbe stata chiamata AI simbolica, mentre quella dedicata alle reti neurali citava le ricerche di McCulloch e Pitts, di Uttley e di due degli organizzatori della conferenza, Minsky e Rochester.

I possibili indirizzi di ricerca descritti nella proposta rivelavano la frattura tra i due approcci in maniera ancora più dettagliata. Tre organizzatori su quattro erano dalla parte del *brain modelling*. Shannon, ad esempio, avrebbe voluto dedicarsi a due argomenti: l’applicazione dei concetti della teoria dell’informazione ai calcolatori e ai modelli del cervello; e l’approccio allo studio degli automi basato sul modello cervello-ambiente.

Il primo progetto, riprendendo alcune delle ricerche di von Neumann e dello stesso Shannon condotte in precedenza, si sarebbe soffermato

¹⁵² Edwards 1996: 253.

sull'affidabilità della computazione a partire dall'impiego di elementi inaffidabili: concetto applicabile tanto alle macchine quanto al cervello umano.¹⁵³ Il secondo progetto si sarebbe invece focalizzato sulla creazione di strutture matematiche per la rappresentazione dell'ambiente con cui il cervello si trovava a interagire: ambiente che non avrebbe dovuto essere né ostile né complesso, perché era tramite la simulazione delle attività più semplici che un giorno si sarebbe potuti giungere, gradualmente, a quella delle attività più avanzate dell'intelletto umano. Minsky sperava di continuare le sue ricerche sull'apprendimento iniziate a Harvard e di completare, prima dell'inizio della conferenza, un modello di rete neurale basato sulla procedura del *trail-and-error* in grado di funzionare su un calcolatore. Rochester voleva approfondire le ricerche condotte dal suo gruppo all'IBM sulla simulazione delle reti neurali tramite calcolatori, in modo che esse manifestassero "originalità" quando poste di fronte a un problema nuovo: cosa che sarebbe stata possibile inserendo un certo grado di causalità (*randomness*) nelle procedure di elaborazione. Fu solo McCarthy, come accennato prima, a soffermarsi sull'elaborazione simbolica tramite calcolatori, proponendo di investigare le relazioni tra intelligenza e linguaggio. La proposta di McCarthy assumeva quasi i caratteri di una dichiarazione generale d'intenti, nel suo sottolineare l'inutilità del metodo cibernetico applicato a determinati contesti:

¹⁵³ Il quale, come si è visto, affrontò lo stesso argomento nell'articolo scritto per gli *Automata Studies*.

It seems clear that the direct application of trial and error methods to the relation between sensory data and motor activity will not lead to any very complicated behavior.¹⁵⁴

L'idea di McCarthy era che il metodo del *trial-and-error* andasse applicato a un più alto livello di astrazione rispetto alla semplice elaborazione sensoriale. Sarebbe stato meglio tentare di costruire un linguaggio artificiale, che un calcolatore sarebbe stato in grado di usare, per rappresentare problemi che richiedevano congetture e auto-referenzialità. Questo linguaggio avrebbe dovuto corrispondere all'inglese, nel senso che ogni breve asserzione su un dato argomento avrebbe dovuto avere il suo corrispettivo simbolico, elaborabile dalla macchina.

McCarthy sembrava essere il solo a intuire la differenza strutturale tra calcolatori digitali e cervello: "The control mechanism of the brain is clearly very different from the control mechanism in today's calculators". Sintomo di questa differenza fondamentale era, ad esempio, la maniera in cui si presentava un guasto. McCarthy ipotizzava che il cervello fosse dotato di un meccanismo che controllava l'ordine sequenziale del comportamento, il quale poteva aggiungere un certo grado di casualità in circostanze particolari (come, ad esempio, quando si è ubriachi) per aumentare l'efficienza dei processi immaginativi. Sulla possibilità di riprodurre questi processi tramite delle reti neurali, la sua posizione era diametralmente opposta a quella di Rochester:

¹⁵⁴ McCarthy et al. 1955: 16.

Some work has been done on simulating neuron nets on our automatic calculator. One purpose was to see if it would be thereby possible to introduce randomness in an appropriate fashion. It seems to have turned out that there are too many unknown links between the activity of neurons and problem solving for this approach to work quite yet.¹⁵⁵

Nonostante l'impegno e una certa originalità d'idee, la proposta non convinse del tutto Morison. Nel tardo settembre del 1955 scrisse a McCarthy che la la proposta "is an unusual one and does not fall easily into our program (di ricerca biologica e medica) so I am afraid it will take us a little time before coming to a decision". Non avendo avuto più notizie da Morison, nel novembre di quell'anno McCarthy chiese a Shannon di contattare Weaver, per vedere a che punto fosse l'esame della proposta. Morison rispose a fine novembre, scusandosi con McCarthy per il ritardo e affermando che la fondazione non avrebbe stanziato la cifra richiesta per il seminario (13500 dollari), ma avrebbe potuto dare il suo contributo con una somma più piccola (7500 dollari). Nella lettera, Morison non aveva nessuna remora nell'affermare lo scetticismo che i membri della fondazione provavano circa la possibilità di matematizzare i processi del pensiero, un campo "very challenging on the long run", ma in quel momento abbastanza difficile da inquadrare. Dal momento che si trattava di un progetto rischioso, la Rockefeller Foundation non volle mettere a disposizione l'intera cifra richiesta.

¹⁵⁵ Ivi.: 14.

McCarthy rispose alla lettera di Morison quasi subito, affermando che avrebbe parlato con gli altri organizzatori della questione, anche se in linea teorica non avrebbero dovuto esserci problemi, al massimo solo qualche cambiamento nei piani originali. McCarthy era abbastanza fiducioso che i soldi messi a disposizione dalla Rockefeller Foundation sarebbero bastati, anche perché, stando alle sue parole, alcuni dei ricercatori invitati avevano garantito che, nel caso in cui si fossero effettivamente presentati, sarebbero state le rispettive organizzazioni a coprire i costi.

Durante la fase di “reclutamento” che seguì, McCarthy fece quello che si sarebbe rilevato un viaggio provvidenziale per visitare Allen Newell e Herbert Simon al CIT (Carnegie Institute of Technology) nel febbraio del 1956.

3.3 Herbert Simon e Allen Newell

La formazione di Simon non sembra giustificare, a prima vista, il suo interesse verso l'AI. Entrato alla RAND come consulente nell'estate del 1952 (due anni dopo l'arrivo di Newell), Simon era già allora un affermato economista, oltre che uno studioso di scienze sociali. Nel 1947 pubblicò il libro *Administrative Behavior*, il cui scopo era dimostrare come le organizzazioni potevano essere comprese nei termini dei loro processi decisionali. Si potrebbe dire che il principale ambito di competenza di Simon in quegli anni (la burocrazia) si collocasse all'esatto opposto rispetto

allo studio dell'intelligenza umana. Ma il carattere innovativo del suo lavoro stava proprio nel trovare i punti di contatto tra questi due campi apparentemente antitetici. Simon era affascinato dal processo decisionale delle persone, giungendo a delle conclusioni che si concentravano proprio sul funzionamento della mente umana, in opposizione alle convenzionali teorie economiche del tempo.

In *Administrative Behavior* vi era una prima esposizione di quella che sarebbe diventata la teoria dell'uomo amministrativo dotato di razionalità limitata, analizzata nel primo capitolo. Prima di compiere, ad esempio, un investimento, una compagnia era solita considerare tutte le possibili alternative e scegliere quindi quella che avrebbe comportato il maggiore profitto. Ci si aspettava che lo stesso comportamento fosse assunto dal consumatore, al momento dell'acquisto. Il libro di Simon spiegava come, attraverso la teoria della razionalità limitata (*bounded rationality*), queste teorie di base non sempre funzionassero nella realtà.

Prima di tutto, nessuno vagliava tutte le possibili alternative prima di una scelta. Analizzare ogni alternativa aveva un costo, nel senso più ampio del termine: di tempo, monetario, ecc. Per questa ragione, un processo decisionale consisteva in una scelta tra un numero limitato di opzioni, che rispettava la regola "the fewer, the better". Invece di considerare attentamente alcune implicazioni di tutte le opzioni (come il profitto associato), il burocrate che è in ognuno sceglieva la prima alternativa

disponibile che soddisfaceva un set preesistente di criteri accettati.

Le possibili alternative che si presentavano nel processo decisionale erano in realtà estremamente limitate, ma questa debolezza non risiedeva necessariamente nella mente. Simon scoprì infatti un'altra debolezza che andava oltre i limiti intellettivi del singolo: sia le persone che le organizzazioni avevano difficoltà a trovare soluzioni originali per i problemi. Nelle organizzazioni, questa incapacità si manifestava tramite l'esistenza dei manuali manageriali, o nei libri di regole. Queste osservazioni portarono Simon a ipotizzare che la mente funzionasse principalmente applicando soluzioni approssimative nella risoluzione dei problemi: idea che sarebbe stata alla base della programmazione euristica.

Simon notò anche che i membri di una stessa organizzazione tendono a perseguire scopi minori (*subgoals*) rispetto a quelli generali. Uno staff incaricato di realizzare uno spot pubblicitario per l'azienda tenderà, presumibilmente, a realizzare uno spot il più incisivo possibile (scopo minore), indipendentemente dal fatto che ciò andrà o meno a beneficio dei profitti dell'azienda (scopo maggiore). Solo l'intervento dei dirigenti poteva riconciliare questi due scopi. In più, le organizzazioni riuscivano a raggiungere i propri scopi generali solo frammentandoli in una serie di scopi più piccoli, con diversi dipartimenti che lavoravano in maniera coordinata. La strategia dei *goal-subgoal* sarebbe diventata uno dei concetti fondamentali dell'AI.¹⁵⁶ In retrospettiva, il passaggio di Simon

¹⁵⁶ Crevier 1993: 42-43.

dall'economia all'AI sembra quasi naturale.

Simon arrivò alla Conferenza di Dartmouth con il suo collega più giovane, Allen Newell. Laureatosi in fisica a Stanford, Newell ebbe modo di seguire i corsi del matematico George Polya. L'AI deve a Polya il termine “euristico”, che egli coniò per descrivere le regole pratiche che si applicano nei ragionamenti di tutti i giorni. Nel 1945, Polya mostrò l'efficacia di questo metodo nella risoluzione dei problemi nel libro *How to Solve it*.

Dopo essersi laureato a Stanford, Newell trascorse un anno a Princeton per specializzarsi in matematica, non trovandosi altrettanto bene quanto Minsky e McCarthy (vi entrò a far parte contemporaneamente a McCarthy, e mancò Minsky solo di qualche anno). Al contrario di loro, Newell non si riteneva un matematico, e lasciò l'istituto, preferendo lavorare dove avrebbe avuto problemi concreti da risolvere. Era il 1950, e in quel periodo la RAND offriva ai giovani scienziati con una certa predisposizione alla pragmaticità la possibilità di mettersi alla prova.

Qui Newell fu assegnato a un progetto che prevedeva di studiare le interazioni uomo-macchina simulando le dinamiche di un centro di difesa antiaerea. Parte dell'attività consisteva nello stampare delle mappe aeree tramite delle macchine tabulatrici. Durante il suo periodo di consulenza alla RAND, Newell ebbe modo di vedere queste macchine in attività. Oggi potrebbe sembrare una cosa banale, visto che praticamente qualunque personal computer è dotato di un programma di disegno, da Paint fino a

quelli più sofisticati. Ma nei primi anni Cinquanta, quella vista fu per Newell una rivelazione: i punti e i caratteri che formavano la mappa non erano numeri. Newell li interpretava come simboli, e la macchina era in grado di manipolarli. Naturalmente, passare da questa considerazione al convincersi che i calcolatori potessero simulare il pensiero umano richiedeva un salto concettuale non indifferente, ma Newell lo fece senza esitazioni. Da quel momento in poi, lui e Simon iniziarono a instaurare delle conversazioni informali sull'argomento.

L'adesione totale a queste idee, tuttavia, avvenne nel 1954, quando Selfridge, allora ai Lincoln Lab, visitò la RAND per parlare delle sue ricerche su un dispositivo in grado di riconoscere le forme visive; ricerche che avrebbero poi portato alla creazione del Pandemonium.¹⁵⁷ La macchina sarebbe stata in grado di riconoscere le lettere e le figure semplici. Per far ciò, un certo numero di sottoprocessi analizzavano le varie caratteristiche della figura, ottenendo dei valori numerici; questi processi avrebbero poi "votato" per il risultato, elaborando un valore e confrontandolo con un set di regole prestabilito. Dal momento che facevano uso di un gran numero di processi semplici, i quali combinavano dei valori per ottenere un risultato, e soprattutto per la loro capacità di "apprendere", perfezionando i propri processi, il Pandemonium e i suoi predecessori si rifacevano chiaramente alle reti neurali di McCulloch e Pitts. Ma erano anche dei manipolatori di simboli, dal momento che ogni sottoprocessso usava la logica per analizzare

¹⁵⁷ Selfridge 1955; Selfridge 1959.

e catalogare le caratteristiche della figura. Proprio come le prime speculazioni di Turing, i programmi di Selfridge occupavano lo spazio di transizione tra il *brain modelling* e l'elaborazione simbolica dell'informazione.¹⁵⁸

Newell avrebbe ripreso da Selfridge l'idea che un processo complesso avrebbe dovuto essere il risultato dell'interazione di molti sottoprocessi più semplici; la stessa idea che, come visto, aveva sviluppato anche Simon, sebbene riferita a un ambito diverso.

3.4 Il *Logic Theorist*

Con l'aiuto di J. C. Shaw, programmatore della RAND, nell'autunno del 1955 Simon e Newell iniziarono a lavorare a quello che viene generalmente considerato il primo programma di AI: il *Logic Theorist*.

All'inizio, furono avanzate tre ipotesi su ciò che il programma avrebbe potuto effettivamente fare. L'intenzione originale era quella di concentrarsi sugli scacchi o sulla geometria. Ma questi ambiti presentavano dei problemi di tipo percettivo, difficili da sintetizzare in un programma, ma fondamentali negli esseri umani quando affrontano questo tipo di sfide. Fu quindi scelta la logica, "for no deeper reason than that I (Simon) had the two volumes of *Principia* (di Russell e Whitehead) at home". Non che l'intenzione di Newell e Simon fosse quella di dimostrare teoremi. Essi cercavano piuttosto di capire come gli esseri umani, tramite la selezione

¹⁵⁸ Edwards 1996: 250-251.

euristica, operavano la scelta migliore: “That was certainly in the center of my mind: what heuristic would kick out the right theorem to use rather than searching forever”.¹⁵⁹

L’intuizione di Simon, Newell e Shaw era che la dimostrazione dei teoremi poteva essere ridotta a una ricerca selettiva, rappresentata graficamente da una struttura ad albero. Il punto di partenza era l’ipotesi iniziale, su cui le regole della logica permettevano un certo numero di manipolazioni elementari, le quali portavano a versioni leggermente modificate dell’ipotesi. Ogni possibile manipolazione corrispondeva a una diramazione dell’albero. L’applicazione di altre manipolazioni a ogni risultato dava la serie seguente di diramazioni, e così via. Alla fine dell’albero, dopo un numero sconosciuto di manipolazioni, vi era la conclusione desiderata: il problema risiedeva nel trovare il percorso giusto.

Per trovarlo, il *Logic Theorist* usava delle procedure euristiche, ovvero l’applicazione di regole che gli avrebbero permesso di selezionare, tra tutte le possibili diramazioni, quella che più probabilmente lo avrebbe portato all’obiettivo. La ricerca, il comportamento orientato a un scopo, le decisioni soggette a regole: queste erano alcune delle idee principali di Simon sulla razionalità limitata, e adesso venivano incorporate nel primo programma euristico della storia.

Come visto nel primo capitolo, negli anni Cinquanta programmare un calcolatore non era impresa da poco, a causa dei limiti tecnici. Il trio di

¹⁵⁹ Crevier 1993: 44.

ricercatori pensò quindi che sarebbe stato meglio simulare prima il programma “dal vivo”, per poi codificarlo dentro la macchina. Il metodo è descritto da Simon nella sua autobiografia:

Al and I wrote out the rules for the components of the program (subroutines) in English on index cards, and also made up cards for the contents of the memories (the axioms of logic). At the GSIA (Graduate School of Industrial Administration) building on a dark winter evening in January 1956, we assembled my wife and three children together with some graduate students. To each member of the group, we gave one of the cards, so that each person became, in effect, a component of the LT computer program subroutine that performed some special function, or a component of its memory. [...] So we were able to simulate the behavior of LT with a computer constructed of human components. Here was nature imitating art imitating nature.¹⁶⁰

L'effettiva programmazione del *Logic Theorist* su un calcolatore non sarebbe avvenuta prima dell'estate del 1956. Tuttavia, questa prima simulazione parve a Simon abbastanza soddisfacente, e lo convinse della validità dell'intuizione. Sempre nella sua autobiografia, Simon non nascondeva l'entusiasmo provato in quel periodo:

¹⁶⁰ Simon 1991: 206-207

We invented a computer program capable of thinking non-numerically, and thereby solved the venerable mind/body problem, explaining how a system composed of matter can have the properties of mind.¹⁶¹

L'esistenza stessa del *Logic Theorist* comportava delle implicazioni filosofiche che fin da subito non sfuggirono ai suoi creatori. Come a confermare le parole di Simon, il *Logic Theorist* era anche in grado di dimostrare 38 dei 52 dei teoremi formulati nei *Principia*. In più, la dimostrazione di uno dei teoremi (il 2.85) si dimostrò perfino più elegante di quella fornita da Russell e Whitehead: Simon contattò Russell per informarlo del suo successo. Inoltre, il programma riuscì a dimostrare il teorema nonostante fosse privo delle istruzioni specifiche necessarie, dimostrando di poter svolgere compiti non contemplati dalla programmazione. Incoraggiati al giudizio positivo di Russell (“ I am delighted to know that *Principia Mathematica* can now be done by machinery”¹⁶²), Newell e Simon decisero di sottoporre la dimostrazione del teorema a Stephen Kleene, editore del *Journal of Symbolic Logic*, in un articolo in cui il *Logic Theorist* figurava come co-autore. Ma la risposta di Kleene fu di tutt'altro tenore: dal suo punto di vista i metodi dei *Principia* erano ormai talmente obsoleti che una dimostrazione fatta ricorrendo a quel sistema non era di nessun interesse. Kleene ignorò totalmente quale fosse il vero senso della scoperta, ma i creatori del *Logic Theorist* decisero di non

¹⁶¹ Ivi.: 190.

¹⁶² Ivi.: 209.

inferire: “It would be rude to suggest that the difference between Kleene's and Russell's responses was further proof of the difference between learning and wisdom”.¹⁶³

L'effettiva implementazione del *Logic Theorist* su un calcolatore richiese più tempo del previsto perché Newell, Simon e Shaw dovettero prima sviluppare un linguaggio di programmazione abbastanza potente e flessibile. Esso venne chiamato IPL (*Information Processing Language*) e incorporava anche un'altra invenzione del trio, forse più importante dello stesso *Logic Theorist*: l'elaborazione a liste (*list-processing*).

L'IPL era diverso da altri linguaggi di programmazione come, ad esempio, il FORTRAN. Per Newell, Simon e Shaw questi linguaggi non erano soddisfacenti perché non riproducevano due importanti caratteristiche del pensiero. Prima cosa, il pensiero sembrava solo consistere nella costante creazione, cambiamento e distruzione di strutture simboliche interagenti tra loro. Linguaggi come il FORTRAN non erano in grado di riprodurre un processo simile. Essi richiedevano che tutte le strutture di numeri e simboli usati dal programma fossero già definite in precedenza nelle sue istruzioni (*program statement*). A queste istruzioni veniva destinata una parte della memoria del calcolatore, in cui depositare questi simboli fino al momento in cui essi sarebbero stati usati. Ciò rendeva impossibile la creazione di nuove strutture di simboli: avendo, per esempio, delle informazioni pre-

¹⁶³ Ibid.

caricate su singoli oggetti dalle caratteristiche comuni, generare una nuova matrice che li racchiudesse sarebbe stato impensabile per linguaggi come il FORTRAN, per il semplice fatto che non era un'operazione contemplata. Ciò diventava invece possibile con l'IPL, che non riservava spazio di memoria alle sue strutture simboliche prima del loro effettivo utilizzo. La seconda caratteristica del pensiero che Newell, Simon e Shaw volevano simulare era il carattere associativo della memoria umana, tramite il quale ogni idea o ricordo poteva portare ad altri simboli a essi collegati. L'abilità di associare e modificare strutture di simboli fu incorporata nel linguaggio di programmazione tramite la tecnica del *list processing*, che Simon descriveva così:

The basic idea is that, whenever a piece of information is stored in memory, additional information should be stored with it telling where to find the *next* (associated) piece of information. In this way the entire memory could be organized like a long string of beads, but with the individual beads of the string stored in arbitrary locations. "Nextness" was not determined by physical propinquity but by an address, or pointer, stored with each item, showing where the associated item was located.¹⁶⁴

Simon e i suoi colleghi presero l'idea dalle prime *drum machine*, gli antenati dei moderni harddisk a memoria fissa. Per ragioni tecniche che non avevano nulla a che fare con la modellazione del pensiero, le *drum machine* usavano già una tecnica simile al *list-processing*. Newell e Simon avrebbero

¹⁶⁴ Ivi.: 212.

discussero dell'idea durante la Conferenza di Dartmouth, e McCarthy l'avrebbe successivamente riutilizzata per sviluppare un altro linguaggio di programmazione, il LIPS (*List Processor*).

La visita di McCarthy al Caltech risale al febbraio 1956, quando Newell e Simon avevano praticamente finito il *Logic Theorist*. Qui, McCarthy mostrò loro la proposta per il seminario di Dartmouth. Dopo la visita, Simon scrisse a Shannon, ripresentandosi.¹⁶⁵ Simon incluse nella lettera anche una serie di annotazioni in cui criticava le ricerche proposte da Shannon per il seminario. Alla luce del lavoro che stava facendo con Newell, Simon scrisse a Shannon che i suoi due progetti di ricerca (l'utilizzo della teoria dell'informazione per analizzare l'elaborazione con componenti inadeguati e la riproduzione dei comportamenti complessi partendo dalle interazioni con ambienti semplici) erano solo in apparenza distinti, ma in realtà riconducibili a un unico progetto:

Because of the rapid progress we have made in the last two months, we are inclined to reverse the strategy expressed in the last sentence of your proposal – we believe that we can start with some of the most advanced human activities (la dimostrazione dei teoremi) and work back to the 'simplest' (il riconoscimento delle singole sillabe).¹⁶⁶

¹⁶⁵ I due avevano avuto modo di conoscersi indirettamente qualche anno prima, quando Simon mandò a Shannon un rapporto della RAND riguardante un programma per giocare a scacchi, ma probabilmente Simon temeva che Shannon non se ne ricordasse.

¹⁶⁶ 21 febbraio 1956.

Shannon non rispose subito alla lettera di sfida di Simon. Due mesi dopo, ad aprile inoltrato, Shannon scrisse a Simon, ammettendo di non aver condotto molte ricerche sull'apprendimento meccanico negli ultimi due anni. Shannon aveva fatto parte della commissione di dottorato di un giovane studente, Trenchard More, al MIT, dove era da poco diventato professore invitato. Shannon scrisse che More stava sviluppando un programma simile a quello di cui parlava Simon, capace di provare teoremi del calcolo proposizionale. Forse per allentare i toni (o forse, al contrario, per rilanciare la sfida), Shannon accluse alla lettera la presentazione del *Symposium on Information Theory*, che si sarebbe tenuto al MIT nel settembre 1956, suggerendo di presentarvi un contributo. "Perhaps", scriveva Shannon, "some of the work you or Newell have been doing would be suitable for a presentation".

Nonostante il tono sempre in bilico tra sfida e cortesia, l'invito fece piacere a Newell e Simon. Al contrario, l'idea che un semplice studente avesse potuto raggiungere un risultato simile al loro sembrò infastidirli molto. Newell avrebbe citato questo episodio in un'intervista di trent'anni dopo, non senza un certo rancore nonostante il tempo passato:

So we have a letter. And in this letter there is some stuff about a guy by the name of Trench More. Trench was a student of Shannon's, and in fact Shannon thought that what Trench was doing was infinitely further ahead than what we were doing.

So there's this funny little exchange and counter exchange and so forth that relates to what Trench is doing and so forth.¹⁶⁷

3.5 Prima, durante e dopo Dartmouth

Nel maggio del 1956, McCarthy scrisse di nuovo a Morison, affermando che i preparativi per il seminario erano ormai quasi completati. Questo avrebbe dovuto tenersi dal 19 giugno al 17 agosto, al Dartmouth College. McCarthy elencava coloro che sarebbero stati presenti per l'intera durata del seminario: egli stesso, Minsky, MacKay, Solomonoff (che in quel periodo, come visto, era collega di Minsky al MIT), John Holland; Jacob Bigelow. Shannon, Rochester e Selfridge sarebbero stati presenti durante le prime due settimane e nelle due finali, mentre Newell e Simon avrebbero partecipato solo alle due settimane iniziali. McCarthy affermava che il numero dei partecipanti era aumentato rispetto ai piani originali perché l'IBM e i Lincoln Lab volevano che vi fosse anche qualcuno dei loro ricercatori, coprendo totalmente le spese.¹⁶⁸ La speranza di McCarthy era che la conferenza si concentrasse sul tipo di ricerche da lui esposte nella proposta:

¹⁶⁷ Norberg 1991: 12.

¹⁶⁸ E' probabile che il numero iniziale di 10 partecipanti fosse stato in un primo momento ridimensionato per sopperire alla mancanza di fondi.

At present it looks as though we shall concentrate on a problem of devising a way of programming a calculator to form concepts and to form generalizations. This, of course, is subject to change when the group gets together.¹⁶⁹

Ma la conferenza non andò come previsto. Diversi nomi indicati da McCarthy non vi presenziarono. MacKay, aggiunto probabilmente su richiesta della Rockefeller Foundation (la quale pagò il suo viaggio dall'Inghilterra tramite un fondo diverso da quello destinato alla conferenza) si ritirò all'ultimo momento. Decise di stare con la moglie, sul punto di partorire il loro primogenito. Questo inconveniente aveva come conseguenza l'eliminazione dal seminario di uno dei più ferventi sostenitori del *brain modelling*.

Holland fece parte del gruppo di Rochester all'IBM durante l'estate del 1955, nello stesso periodo in cui vi lavorò McCarthy. Da quell'anno, insieme ad alcuni colleghi della University of Michigan, Holland iniziò a organizzare dei corsi estivi incentrati sulla teoria degli automi, cui parteciparono anche Simon, Newell, Minsky e McCarthy. Ma sebbene il suo nome figurasse tra quello dei partecipanti alla Conferenza di Dartmouth, non vi prese mai parte:

At that time I had heavy commitments at Michigan. I can't remember why I didn't go, because I planned to. But I did not and that was my great loss, because that was a very important meeting.¹⁷⁰

¹⁶⁹ 25 maggio 1956.

Forse più inaspettata fu l'assenza di Shannon, che probabilmente mandò Trench More in sua rappresentanza, anche per continuare la discussione iniziata con Newell e Simon. Bigelow, pioniere della cibernetica insieme a Wiener e Rosenblueth, non viene citato in nessuno dei resoconti postumi del seminario, e tuttora non si conoscono i motivi della sua assenza. Di contro, subentrarono due nuovi partecipanti inizialmente non previsti: Herbert Gelernter e Arthur Samuel.

Alla fine, i 10 partecipanti ufficiali furono Gelernter, McCarthy, Minsky, More, Newell, Rochester, Samuel, Selfridge, Simon e Solomonoff. Altri vennero solo per brevi visite, come Alex Bernstein, programmatore all'IBM, invitato a parlare del programma di scacchi a cui stava lavorando. Il suo lavoro era già noto a Shannon, Rochester e Samuel, che in quel periodo stava egli stesso lavorando al primo programma per giocare a dama.¹⁷¹

Bernard Widrow (che superò il suo esame di dottorato l'estate di quell'anno) sentì parlare della conferenza da un collega del MIT: "He explained to me a little bit about artificial intelligence", e ne restò incuriosito. I due guidarono da Cambridge fino a Dartmouth di propria iniziativa, senza nessun invito, ma venendo comunque ben accolti dal gruppo ufficiale. Stando alle parole di Widrow, il gruppo includeva anche W.A. Clark e B.G. Farley dai Lincoln Labs, anche loro lì probabilmente come visitatori. Widrow era "fascinated by the things I heard (a Dartmouth)

¹⁷⁰ Husbands 2008a: 389.

¹⁷¹ McCorduck 1997: 113.

about artificial intelligence”, e la cosa avrebbe influenzato le sue ricerche future sulle reti neurali adattive.¹⁷²

Uno dei motivi per cui vi è incertezza su chi fosse effettivamente presente all’evento risiede nel fatto che McCarthy, nonostante avesse promesso di tenere Morison sempre aggiornato, non scrisse mai dei veri resoconti per la Rockefeller Foundation. A essa spedì solo un lungo documento scritto da Solomonoff sull’inferenza induttiva, promettendo che una sessione dedicata alla conferenza (ovvero al suo svolgersi e ai suoi risultati) si sarebbe tenuta durante il *Symposium on Information Theory*, previsto per il settembre seguente, dove “all of speakers (Simon, Newell, Rochester, Minsky, Solomonoff e McCarthy) will be participants in our summer project”. Ma saranno solo gli articoli di Newell-Simon (*The Logic Theory Machine*) e di Rochester sulle reti neurali (*Tests on a Cell Assembly Theory of the Action of the Brain, Using a Large Digital Computer*) ad apparire negli atti del simposio, probabilmente perché gli altri ricercatori non riuscirono a rispettare le scadenze di consegna.

Simon ricordava alcune “difficili contrattazioni” riguardanti il simposio. In teoria, sarebbe dovuto essere McCarthy a esporre le ricerche condotte a Dartmouth, ma Newell e Simon si opposero duramente:

Beh, noi pensavamo a come impedire che questo accadesse, e così il povero Walter Rosenblith (collega di Wiener al MIT), che presumibilmente avrebbe

¹⁷² Anderson, Rosenfeld 1998: 48-49.

presieduto la sessione, passeggiò con noi per il campus del MIT, gironzolando giù per Mem Drive, negoziando la cosa. Non eravamo affatto soddisfatti di John (McCarthy), e Rosenblith cercava di porsi in un angolo neutrale, mentre noi pensavamo che non ci fosse nessun angolo neutrale.¹⁷³

Alla fine, Newell e Simon rimasero soddisfatti di quello che accadde. McCarthy tenne un discorso generale e poi Newell e Simon presentarono dettagliatamente il loro lavoro, in quello che fu il primo annuncio pubblico del *Logic Theorist*. McCarthy avrebbe ammesso, anni dopo, che Newell e Simon “avvertivano, forse del tutto giustamente, che la situazione era anomala, dato che la relazione sul congresso era tenuta da gente che in realtà non aveva fatto nulla, mentre essi avevano fatto qualcosa”.¹⁷⁴

Shannon aveva ormai del tutto abbandonato le ricerche sugli automi, presentando un articolo sulla codificazione dei segnali disturbati.¹⁷⁵

La Rockefeller Foundation non era sicura se considerare la Conferenza di Dartmouth un successo o meno. Probabilmente apprezzarono la responsabilità fiscale mostrata dagli organizzatori, i quali restituirono circa 1700 dollari dei 7500 ricevuti come fondi non spesi; di certo fu molto meno gradita la mancanza di un rapporto dettagliato. Selfridge discusse con Morison, a metà 1957, riguardo la possibilità di finanziare il viaggio di alcuni ricercatori americani affinché presenziassero a una conferenza sull'AI al National Physical Laboratory, a Teddington, di cui Selfridge e

¹⁷³ McCorduck 1997: 128.

¹⁷⁴ Ivi.: 127.

¹⁷⁵ Shannon 1956: 221-238.

Minsky sarebbero stati gli organizzatori americani. Morison era dubbioso: scrisse sul suo diario di aver detto a Selfridge “there has been a little feeling here that we were not entirely well informed to what went on at the conference held in Hanover last summer”. Nonostante tutto Morison si congedò dal colloquio con “an excellent impression of S. (Selfridge), and there is no doubt about the qualifications of the most of the other people who would be invited to the conference (di Teddington)”.

La lista dei partecipanti includeva Minsky e McCarthy, ma anche cibernetici “puri” come Pitts, Ashby e Uttley. Shannon e MacKay erano nella commissione consultiva. Morison sembrava provare adesso meno scetticismo per l’AI, ammettendo che “the field is an important one and is developing very rapidly”.¹⁷⁶ Ma un mese dopo, nel giugno del 1957, la fondazione disse che aveva bisogno di maggiori informazioni prima di stanziare dei finanziamenti.

Tenuta a Teddington nel tardo 1958, la conferenza sulla *Mechanisation of Thought Process* fu una pietra miliare nella storia dell’AI. Organizzata da Uttley in Inghilterra e da Selfridge negli Stati Uniti, il suo scopo era quello di “to bring together scientists studying artificial thinking, character and pattern recognition, learning, mechanical language translation, biology, automatic programming, industrial planning and clerical mechanization”. Essa infatti riunì i più importanti ricercatori impegnati nei principali approcci all’intelligenza meccanica del tempo. Vennero presentati alcuni

¹⁷⁶ 8 maggio 1957.

scritti sull'AI simbolica ad opera di Minsky, McCarthy e Selfridge; sulle reti neurali da parte di McCulloch, Uttley, Frank Rosenblatt; e da Ashby, MacKay e Gordon Pask sui sistemi a reti non-neurali auto-organizzate.

Gli atti di questa conferenza contengono alcuni degli articoli che si sarebbero rivelati maggiormente influenti per il futuro dell'AI. Lo scritto di Minsky, *Some Methods of Artificial Intelligence and Heuristic Programming* (lo stesso in cui affermava la sua conversione totale al metodo simbolico), era l'ultima versione di un articolo provvisorio cui stava lavorando fin dai tempi della Conferenza di Dartmouth. Il documento descriveva i vari metodi che erano (e avrebbero potuto essere) utilizzati nella programmazione euristica. Esso si occupava anche dei metodi tramite cui sarebbe stato possibile il riconoscimento delle forme, l'apprendimento e la pianificazione tramite calcolatori. La versione finale sarebbe stata pubblicata nel 1961, con il titolo *Steps Toward Artificial Intelligence*.

La speranza di McCarthy, invece, restava quella di sviluppare un linguaggio artificiale. E' così che riassumeva il suo articolo, *Programs with Common Sense*:

This paper will discuss programs to manipulate in a suitable formal language (most likely a part of the predicate calculus) common instrumental statements. The basic program will draw immediate conclusions from a list of premises.

These conclusions will be either declarative or imperative sentences. When an imperative sentence is deduced, the program takes a corresponding action.¹⁷⁷

McCarthy suggeriva che i fatti che un ipotetico programma d'intelligenza artificiale (che egli chiamava *Advice taker*) avrebbe dovuto elaborare potevano essere rappresentati come espressioni formulate in un linguaggio matematico (e *computer-friendly*) della logica di primo ordine (*first-order logic*). Ad esempio, fatti come "I am at my desk" e "My desk is at home" sarebbero stati rappresentati così: at (I, desk) e at (desk, home). Questi fatti, insieme a espressioni che avrebbero rappresentato informazioni su come spostarsi da un luogo all'altro (a esempio guidando o camminando), avrebbero potuto essere utilizzati dal proposto (ma non ancora effettivamente programmato) *Advice taker* per raggiungere determinati obiettivi. Nel caso si volesse, ad esempio, raggiungere l'aeroporto, il processo di elaborazione dell'*Advice taker* avrebbe prodotto un'espressione logica imperativa riguardante la necessità di camminare fino alla macchina e poi guidare fino all'aeroporto.

McCarthy avrebbe successivamente esteso queste idee in una raccolta di memorandum, e alcune delle sue idee per l'*Advice taker* saranno effettivamente realizzate da uno studente di Stanford: C. Cordell Green.

Con Newell, Simon e McCarthy l'AI degli esordi era legata

¹⁷⁷ McCarthy 1956: 77.

indissolubilmente alla logica, ai suoi processi e alle sue regole: negli anni, le cose sarebbero cambiate.

3.6 Un congresso controverso

Il *Dartmouth Summer Research Project on Artificial Intelligence* è generalmente considerato da storici e informatici il luogo di nascita dell'AI come campo di ricerca scientifica. Secondo Pamela McCorduck, sono due gli eventi che giustificano questa scelta: il debutto ufficiale di un nuovo paradigma scientifico, basato sull'elaborazione simbolica delle informazioni e incarnato dal *Logic Theorist* di Newell e Simon, che avrebbe dominato la ricerca sull'AI per tutto il decennio successivo; e la nascita di “modelli sociali” tra i partecipanti, i quali avrebbero successivamente creato i loro primi programmi al MIT, alla Stanford University, alla Carnegie Mellon University e al Stanford Research Institute. In seguito alla conferenza questi quattro istituti sarebbero diventati in assoluto i più importanti e influenti nel nuovo ambito di ricerca, oltre che i più finanziati dall'ARPA (Advance Research Projects Agency) del Dipartimento della Difesa, e tali sarebbero rimasti per tutto il ventennio seguente: “Forse è nella natura dei nuovi campi scientifici l'essere dominati dai loro fondatori, e dalle istituzioni in cui i loro fondatori sono di casa”.¹⁷⁸

I partecipanti hanno espresso, nel corso degli anni, opinioni divergenti sul significato da attribuire alla conferenza. Selfridge ricordava di aver avuto la

¹⁷⁸ McCorduck 1997: 130.

sensazione che fosse stato concepito qualcosa d'importante, anche se nessuna delle promesse avanzate dalla conferenza venne mantenuta nei tempi previsti.¹⁷⁹ Newell la definì “a piece of AI mythology”, dal momento che creò un senso d'identità nella neonata comunità scientifica.¹⁸⁰

La conferenza fu per certi versi inconcludente e profondamente deludente per il suo organizzatore principale, John McCarthy. Molti degli invitati vi presenziarono per periodi di tempo diversi, rendendo quasi impossibili gli incontri regolari. Simon ricordava che:

They were going to have a kind of floating crap game all summer with people sitting with each other, thinking and so on. And (lui e Newell) were busy programming the Logic Theorist, so we agreed we'd spend a week at Dartmouth.¹⁸¹

Non emerse neanche alcun consenso unanime su cosa fosse effettivamente il nuovo campo di ricerca e quali direzioni avrebbe dovuto prendere. Anche dopo la conferenza, molti dei partecipanti persistettero con i loro approcci e le loro idee. E' sempre Simon a sottolineare, non senza un certo compiacimento, il generale sentimento di orgoglio e testardaggine che li caratterizzò durante il congresso:

It was going off into different directions. They didn't want to hear from us, and we didn't want to hear from them: we had something to show them! In a way, it

¹⁷⁹ Ivi.: 114.

¹⁸⁰ Norberg 1991: 116.

¹⁸¹ Crevier 1993: 49.

was ironic because we already had done the first example of what they were after; and second, they didn't pay much attention to it.¹⁸²

Minsky ricordava di aver provato un falso senso di realizzazione. Ma contrariamente a quanto credevano i partecipanti, la loro comprensione delle teorie riguardanti la manipolazione simbolica era ancora piuttosto limitata. In più, le loro ricerche non attirarono l'attenzione generale della comunità scientifica nella maniera in cui speravano: i gruppi di ricerca continuarono a essere piuttosto ristretti anche negli anni immediatamente successivi, e quasi nessuno dei programmi creati in quel periodo sembrava funzionare davvero. Ciò si tradusse in un generale senso di sfiducia sull'effettiva praticabilità di quell'approccio. Ma il punto di forza delle prime ricerche sull'AI stava anche in questo:

Most people thought AI was impossible. That was the pleasant part of it: if you got an idea, you didn't have to worry about publishing it the same week!¹⁸³

Eppure, l'importanza storica della conferenza è indubbia. Molti s'incontrarono lì per la prima volta, quasi cristallizzando il gruppo. In più, per tutto il ventennio successivo, i principali successi nel campo dell'AI sarebbero stati ottenuti da questi stessi scienziati o dai loro studenti. Dopo Dartmouth, l'AI era ormai, nel bene e nel male, un campo di ricerca intellettuale a tutti gli effetti, per quanto controverso. Per certi aspetti non

¹⁸² Ibid.

¹⁸³ Ibid.

era più unificato di quanto non lo fosse prima del 1956 ma, forse per il continuo scambio d'idee iniziato a Dartmouth, l'AI, da quel momento in poi, iniziò a progredire molto più rapidamente. Già nell'immediato post-Dartmouth fu possibile vedere i risultati di questo processo. Nel 1957, ad esempio, Newell e Simon completarono il *General Problem Solver*, diretta evoluzione del *Logic Theorist*:

The purpose of this paper is not to relate the program to human behavior, but to describe its main characteristics and to assess its capacities as a problem-solving mechanism. The paper will present enough theoretical discussion of problem-solving activity so that the program can be seen as an attempt to advance our basic knowledge of intellectual activity.¹⁸⁴

Non si trattava più di ingegnerizzare i processi cognitivi umani per renderli compatibili con un dispositivo elettronico: al contrario, si trattava di trovare un linguaggio computerizzato adatto a rappresentare i processi cognitivi umani. Dopo Dartmouth, Newell e Simon si resero conto che il comportamento del *Logic Theorist* non rispecchiava pienamente quello umano. La differenza tra il *Logic Theorist* e il *General Problem Solver* risiedeva nel fatto che quest'ultimo era dotato di un'organizzazione e di un set di procedure euristiche che non facevano riferimento a nessun compito particolare (*task-independent*): da qui il *general* nel nome. Alla macchina veniva fornita una tabella differenziale (*difference table*), contenente tutte le

¹⁸⁴ Newell, Shaw, Simon 1958.

informazioni relative al problema (scopo, posizione spaziale degli oggetti, informazioni su di essi, ecc.): sarebbe stata essa stessa a risolvere il problema tramite un'analisi mezzo-fine (*means-ends*), trovando le azioni necessarie e l'ordine in cui andavano compiute. Newell e Simon erano così sicuri della validità del loro approccio da fare anche quattro importanti previsioni per il futuro: musica di un certo valore estetico sarebbe stata composta da un calcolatore; la maggior parte delle teorie psicologiche avrebbero assunto la forma di programmi computerizzati; un calcolatore avrebbe trovato e provato un importante e sconosciuto teorema matematico; un calcolatore sarebbe stato campione mondiale di scacchi.

Nel 1958, McCarthy sviluppò il linguaggio di programmazione LISP (*LIS*t *Processign*), famoso per il suo alto livello di astrazione simbolica. Il LISP manipola liste di simboli di ogni tipo, e liste di liste; i programmi scritti tramite LISP sono essi stessi liste. Questo linguaggio è altamente ricorsivo, e ciò significa che la definizione di un termine può includere quello stesso termine, e che i processi del programma possono richiamare se stessi. Uno dei vantaggi principali del LISP era la capacità di poter manipolare termini dotati di molte interrelazioni complesse, come le parole del linguaggio naturale.¹⁸⁵ Quello stesso anno, McCarthy e Minsky fondarono l'Artificial Intelligence Group nel Research Laboratory of Electronics al MIT; un anno dopo Selfridge avrebbe ultimato il Pandemonium.

Ma per gli sviluppi successivi, furono i finanziamenti dell'ARPA a giocare

¹⁸⁵ Edwards 1996: 354.

un ruolo determinante. Ricordando le parole di Weaver, un'altra delle particolarità della Conferenza di Dartmouth fu quella di essere un progetto "personale", nel senso che non vi era nessuna grossa istituzione dietro: da lì la necessità dei finanziamenti. Ciò alla fine non si rivelò del tutto vero: sebbene lo sponsor principale fosse senza dubbio la Rockefeller Foundation, anche l'Office of Naval Research finì per essere coinvolto. Negli anni Cinquanta, la politica dell'ONR era quella di puntare sulle tecnologie computerizzate per le decisioni di supporto, intercettando qualunque progetto utile allo scopo nelle università. Fu per questo che decise di finanziare la conferenza, sebbene Martin Denicoff, uno dei suoi dirigenti, condividesse lo stesso scetticismo manifestato da Morison:

The Conference was so novel at the time that no one really knew what would come of it. In fact, I questioned that the participants had a good insight into what was going to happen.¹⁸⁶

I dubbi erano ben fondati per quanto riguardava la mancanza di organizzazione, ma non per quanto riguardava i futuri progressi. Subito dopo la conferenza, l'ONR stipulò dei contratti a lungo termine con alcuni scienziati del Carnegie Tech (inclusi Newell e Simon), "essentially to explore new approaches to decision making with all of these people involved".¹⁸⁷

Ma anche allora l'AI restava solo uno dei tanti possibili approcci alla

¹⁸⁶ Denicoff 1988: 271.

¹⁸⁷ Ivi.: 272.

relazione uomo-macchina: una minuscola parte di un programma di ricerca molto più grande, e tale probabilmente sarebbe rimasta se non fosse stato per il contributo dell'ARPA.

Fondata nel 1957 dall'amministrazione Eisenhower in seguito al lancio dello Sputnik, l'ARPA era inizialmente un'agenzia spaziale provvisoria, un grosso contenitore in cui riversare tutti i programmi spaziali (militari e civili) in attesa che il governo creasse un'istituzione più duratura. Quando, nel 1960, all'inizio dell'amministrazione Kennedy, i programmi spaziali vennero trasferiti alla NASA, l'ARPA continuò comunque a disporre di un gran numero di ricerche sperimentali già avviate, scientifiche e tecnologiche.

Il ruolo dell'ARPA post-NASA divenne quello di creare tecnologie volte alla difesa del territorio, di un tipo che i servizi militari e gli altri dipartimenti, per vari motivi, non potevano o non volevano sviluppare: perché i rischi erano troppo alti, perché non rientravano nei loro piani a breve o lungo termine, ecc. Nel suo unire ricerca sperimentale e sicurezza nazionale, l'ARPA può per certi versi essere vista come corrispettivo di quello che fu l'OSRD durante la Seconda guerra mondiale. Non dovendo rispondere a nessuno se non al Dipartimento della Difesa, e potendo ignorare le tendenze di ricerca delle altre agenzie, l'ARPA poté sviluppare progetti allora considerati molto rischiosi, ma la cui utilità sarebbe andata oltre la necessità militare contingente, nel tentativo di evitare le “sorpresa

tecnologiche” da parte degli altri paesi.

Viste le sue finalità militari, le ricerche sponsorizzate dall'ARPA non erano soggette al *peer review*: lo stesso piccolo direttorio di scienziati e di ingegneri sceglieva quali ricerche finanziarie in base al proprio giudizio professionale, senza la minima supervisione. L'ARPA decise quindi di concentrare le proprie risorse finanziando un piccolo numero di centri d'eccellenza, arricchendo quelli che sarebbero stati i quattro centri principali per il futuro dell'AI: la Carnegie-Mellon University, dove lavoravano Newell e Simon; il MIT, dove lavorava Minsky; Stanford, dove lavorava McCarthy; e lo Stanford Research Institute, frequentato dagli allievi di alcuni dei partecipanti della conferenza.

Nonostante il coinvolgimento di così tante menti brillanti sia stato senza dubbio determinante nello sviluppo dell'AI, la volontà delle agenzie di finanziare una piccola élite d'istituti non è un fattore da sottovalutare.

La Conferenza di Dartmouth si tenne tre anni dopo l'ultima Conferenza Macy sulla cibernetica. La cibernetica nacque inizialmente quasi come un enorme ricettacolo in cui far confluire tutte le ricerche che facevano capo all'analogia uomo-macchina e di cui l'AI, all'inizio, era solo una delle tante specificazioni possibili, forse tra le più controverse, ma sicuramente non tra le più popolari.

I due movimenti sembrarono coesistere almeno fino alla metà degli anni Sessanta, quando la contraddizione di fondo divenne chiara ai più. L'AI era

ormai un campo di ricerca indipendente, nato dalla cibernetica ma da essa completamente autonomo. Questa frattura venne illustrata da Minsky nel 1968, nel suo *Semantic Information Processing*:

While work on artificial intelligence draws upon methods from other fields, this is not a significantly interdisciplinary area: it has its own concepts, techniques, and jargon, and these are slowly growing to form an intricate, organized specialty.¹⁸⁸

Per Minsky, neanche le ricerche di Newell e Simon andavano inquadrare sotto la categoria di quella che egli intendeva come AI:

Cybernetics divided, in my view, into three chief avenues. The first was the continuation of the search for simple basic principles (i sistemi minimali auto-organizzati) - most clearly exemplified by the precise analyses of Ashby. The second important avenue was an attempt to build working models of human behavior incorporating, or developing as needed, specific psychological theories. Work in this area (simulation of human thought) has focused rather sharply at the Carnegie Institute of Technology, in the group led by Newell and Simon. The third approach, the one we call Artificial Intelligence, was an attempt to build intelligent machines without any prejudice toward making the system simple, biological, or humanoid. Much of the earlier work on artificial intelligence was done by the group at MIT.¹⁸⁹

Un anno dopo, Minsky avrebbe denunciato i limiti delle reti neurali nel libro (scritto insieme a Seymour Papert) *Perceptrons: an introduction to*

¹⁸⁸ Minsky 1968: 6.

¹⁸⁹ Ivi.: 7-8.

computational geometry.

Nei primi anni Sessanta molte università conducevano ancora ricerche sulle reti neurali, e alcune delle più importanti tra esse facevano capo a Frank Rosenblatt, il quale costruì quella che può essere definita la macchina più famosa di quel periodo: il *Perceptron*.¹⁹⁰

Il *Perceptron* di Rosenblatt era una rete neurale capace di riconoscere le forme.¹⁹¹ Per migliorarne le prestazioni durante le opere di classificazione, le connessioni modificabili della rete neurale potevano essere regolate in base a determinate regole (o algoritmi di apprendimento). Nel 1960, il team di ricerca guidato da Rosenblatt, Widrow e Hoff, sviluppò due importanti algoritmi di apprendimento per le reti neurali a un solo layer. Rosenblatt dimostrò che, se un *Perceptron* fosse stato fisicamente in grado di eseguire un compito di classificazione (ovvero se i suoi parametri fossero stati in grado di incarnare questo compito), allora a esso si sarebbe potuto “insegnare” questo compito tramite un numero finito di cicli di allenamento (*training cycles*), i quali comprendevano la presentazione di un modello (*pattern*), l’osservazione dell’output dato dalla macchina e il miglioramento delle connessioni seguendo un algoritmo.¹⁹²

Il lavoro Minsky e Papert cercava di dimostrare come ulteriori progressi nel campo delle reti neurali fossero impossibili, e come questo approccio

¹⁹⁰ Olazarin 1996.

¹⁹¹ Anche se solo le connessioni che collegavano le unità di associazione agli output erano regolabili.

¹⁹² Rosenblatt 1960.: 111.

mancasse di valore scientifico e pratico. Essi analizzavano un *Perceptron* dotato di un numero ristretto di connessioni modificabili. In breve, Minsky e Papert dimostrarono che i cicli richiesti da un *Perceptron* per calcolare la parità (*parity*) (cioè se il numero di input attivati in una retina fosse pari o dispari) e la concatenazione (*connectedness*) (ovvero stabilire se un insieme di punti attivati nella retina appartenesse a uno stesso oggetto – cioè se fossero collegati tra di loro) non erano finiti, ma aumentavano con la dimensione degli input della retina.

Minsky e Papert non si limitarono a sottolineare questi limiti, ma illustravano anche come tutti questi problemi avrebbero potuto facilmente essere aggirati ricorrendo a un calcolatore:

Many of the theorems show that perceptrons cannot recognize certain kinds of patterns. Does this mean that it will be hard to build machines to recognize those patterns? No. All the patterns we have discussed can be handled by quite simple algorithms for general-purpose computers.¹⁹³

Harry Collins faceva notare come attacchi del genere fossero abbastanza inusuali nel mondo della scienza, dove in genere si accettano i risultati di un'area di ricerca non direttamente collegata alla propria, per quanto diametralmente opposti: se succedeva, era perché la controversia (in questo caso *brain modelling* contro AI simbolica) aveva assunto una certa importanza storica, finendo con il coinvolgere i principali “nuclei” di

¹⁹³ Minsky, Papert 1969: 72.

scienziati interessati alla questione, rappresentati in questo caso da Minsky e Rosenbluett.¹⁹⁴

Negli anni che seguirono, i maggiori finanziamenti e l'accresciuto interesse nei confronti dell'AI andarono del tutto a sfavore della cibernetica, che si vide privare delle sue principali forme di sostentamento. Le ricerche sulle reti neurali, sui sistemi auto-organizzati e sull'apprendimento meccanico vennero di fatto oscurate per i decenni successivi, spostando gli interessi della disciplina verso il campo delle scienze sociali e della terapia.¹⁹⁵

Già nel 1958, MacKay iniziò a essere prudente nell'utilizzare il termine "cibernetica", per via del dilagare di gruppi "devianti" che, per quanto ne riprendessero il nome, si occupavano di ricerche che riguardavano, ad esempio, la creazione di organismi artificiali, e che quindi avevano poco a che spartire con la cibernetica degli inizi.¹⁹⁶

In una lettera indirizzata a Heinz von Foerster, MacKay esprimeva la delusione per l'adesione di quest'ultimo all'ARTORG (Artificial Organism Research Group):

It's for just this kind of reason that folk such as Gabor, Uttley, Cherry & I are chary of using the word 'Cybernetics' nowadays, and I do hope that the work of someone of your calibre won't lose some of the attention it deserves by this new

¹⁹⁴ Collins 1981: 6-10.

¹⁹⁵ Cariani 2009:89.

¹⁹⁶ Kline 2009: 353.

connection. (I assume that you have seen the earlier ‘Artorg’ sheets? Even Warren (McCulloch) seemed a bit shocked by the one I showed him!).¹⁹⁷

Bisognerà aspettare il grande *revival* delle reti neurali, avvenuto durante gli anni Ottanta, per far sì che il rinnovato interesse nei confronti di questa scienza permettesse agli storici di dare la giusta collocazione alla cibernetica nei primi anni dell’AI, e dell’AI nella storia della prima cibernetica.¹⁹⁸

¹⁹⁷ Tratto dai *Von Foerster Papers*, box 7, 25 maggio 1959.

¹⁹⁸ Cordeschi 2002; Husband, Holland, Wheeler 2008.

4. L'AI in Italia negli anni Cinquanta-Sessanta

Non meno che altrove, anche in Italia la cibernetica fu fin dalle origini oggetto dell'interesse di un gruppo eterogeneo di studiosi, con diversi centri di ricerca attivi nel periodo del suo maggiore sviluppo, all'incirca tra i primi anni Cinquanta (cioè poco dopo la pubblicazione del libro di Wiener *Cybernetics: Or Control and Communication in the Animal and the Machine*) e i primi anni Settanta del Novecento.¹⁹⁹

Ma se l'Italia si dimostrò particolarmente ricettiva nel cogliere e far propri i contenuti della cibernetica, nello stesso periodo l'AI fu invece oggetto di studio di un numero estremamente limitato di ricercatori, i quali adottarono nei suoi confronti un punto di vista spesso molto critico, seppur non privo di obiezioni fondate e in anticipo sui tempi.

4.1 *Adamo II*

Correva l'anno 1953, e sfogliando le pagine di *Civiltà delle Macchine*, la rivista aziendale della Finmeccanica fondata dal poeta-ingegnere Leonardo Sinisgalli (e da lui diretta fino al 1958), era possibile imbattersi nel seguente “semaforo”:

COLLOQUIO POSSIBILE – Secondo quanto scrive A. M. Turing sulla maggiore rivista inglese di filosofia, *Mind*, le difficoltà che incontrerebbe una macchina per

¹⁹⁹ Marchis 2013.

dare delle risposte simili a quelle tradotte qui sotto sono puramente tecniche. C'è qualcuno che le trovi sufficienti per dire che la macchina “pensa” e “ragiona”?²⁰⁰

Seguiva la traduzione italiana dell'ipotetico dialogo tra esaminatore e macchina immaginato da Turing nell'articolo *Computing Machinery and Intelligence* del 1950. “Semaforo” era una rubrica redazionale che riportava brevi notizie, curiosità e precisazioni, collocata nelle ultime pagine di ogni fascicolo della rivista, famosa soprattutto per aver proposto articoli e lettere di rinomati autori con lo scopo di creare un confronto fra scrittori, poeti e pittori con il mondo dell'industria e della tecnologia, negli anni che portarono al miracolo economico italiano.²⁰¹

A soli tre anni di distanza dalla loro pubblicazione originale, le ricerche di Turing trovavano eco anche all'interno di una rivista aziendale italiana, e proprio in relazione a uno dei loro aspetti più delicati e controversi, cioè la possibilità che una macchina potesse arrivare a livelli di ragionamento intelligente paragonabili a quelli umani.²⁰² La cosa non meraviglia del tutto, dato che Sinisgalli fu uno dei maggiori sostenitori del progetto denominato (da lui stesso) *Adamo II*, il primo automa creato dalla scuola cibernetica italiana, guidata dal filosofo e linguista Silvio Ceccato.

C'è una cosa che colpisce nel riassunto di Sinisgalli: l'aver del tutto

²⁰⁰ Sinisgalli 1953b: 79.

²⁰¹ <http://matematica.unibocconi.it/articoli/civilt%C3%A0-delle-macchine-leonardo-sinisgalli-e-alan-turing#1>

²⁰² Una citazione (indiretta) al gioco dell'imitazione di Turing si trova anche nell'articolo di Bruno de Finetti *Macchine “che pensano” (e che fanno pensare)*, pubblicato nel 1952 sulla rivista *Tecnica ed organizzazione*.

trascurato l'analisi del dialogo tra esseri umani, che nell'articolo di Turing era premessa necessaria per aggirare il quesito “possono le macchine pensare?”. Citando Turing:

Likewise according to this view the only way to know that a man thinks is to be that particular man. It is in fact the solipsist point of view. It may be the most logical view to hold but it makes communication of ideas difficult. A is liable to believe “A thinks but B does not” whilst B believes “B thinks but A does not” instead of arguing continually over this point it is usual to have the polite convention that everyone thinks.²⁰³

Sembra quasi che Sinisgalli abbia voluto bruciare le tappe e andare dritto al punto, in una sorta di monito. Quest'idea viene ulteriormente rafforzata dalla recensione che nel numero precedente di *Civiltà delle macchine* egli aveva fatto del primo *Congresso Italiano di Metodologia*, tenutosi tra il 17 e il 21 dicembre 1952 e organizzato dal Centro di Studi Metodologici di Torino e dal Centro Italiano di Metodologia e Analisi del Linguaggio, che ai tempi pubblicava la rivista *Methodos*. La recensione di Sinisgalli si soffermava in modo particolare sull'ultima relazione del Congresso, quella di Delfino Insolera, dal titolo *Considerazioni sulla tecnica matematica richiesta dalle macchine calcolatrici ad alta velocità*, della quale scriveva:

²⁰³ Turing 1950: 439.

Resta forse irriducibilmente mentale, cioè umana, l'operazione di invenzione dei nuovi algoritmi, cioè in sostanza l'invenzione di nuove operazioni: ma questo vuol dire che la macchina calcolatrice non è un matematico.²⁰⁴

La “macchina” era insomma condannata a replicare automaticamente le procedure nate dall'ingegno umano. Questo avvertimento, unito al “semaforo” pubblicato nel numero successivo e, quasi in contemporanea, all'aiuto offerto a Ceccato, era solo in apparenza contraddittorio. Che Sinisgalli, nonostante un certo pregiudizio (all'epoca piuttosto diffuso) sulla possibilità di creare “macchine pensanti”, fosse comunque interessato a tutti i possibili sviluppi dell'automazione lo confermava lo stesso Ceccato in occasione di un convegno del 1982 in onore dell'intellettuale lucano:

L'idea (per la costruzione di *Adamo II*) m'era venuta in un soggiorno a Londra, nella primavera del 1953. [...] Passando per Roma, qualche tempo dopo, ne parlai con un vecchio amico, Leonardo Sinisgalli, una delle persone più vivaci di ingegno e più disponibili che avessi conosciuto, sin dai tempi della nostra vita goliardica a Milano. L'idea di una macchina “mentale” lo attrasse, forse perché un tanto di magica fantasia non dispiaceva nemmeno a lui. E certo perché non ne vide le implicazioni filosofiche, implicazioni destinate a pesare su ogni sistema religioso ed ideologico, che di valori etici assolutizzati si avvalgono. A dire il vero, nemmeno io gettavo lo sguardo così lontano in direzione pratica. Una macchina è una macchina ed un uomo un uomo. Forse che il calcolatore ha mai disturbato

²⁰⁴ Sinisgalli 1953a: 80.

nessuno? Che chi è tenuto a far calcoli, sicuri ed in fretta, potrebbe farne a meno?
Ma per questo rifiuta di iscriversi ad un partito o credere in qualche divinità?²⁰⁵

Era lo stesso Ceccato ad ammettere come in quel periodo regnasse ancora una certa inconsapevolezza su quali fossero le differenze sostanziali fra una macchina intesa come strumento di pura automazione e una macchina presa come modello delle operazioni mentali; differenza che, se fosse stata chiara a Sinisgalli, aggiunse ironicamente (ma non troppo), lo avrebbe probabilmente portato a fare “un salto indietro, come tante volte ho visto fare in seguito anche dai miei fautori più entusiasti”.²⁰⁶

L’intenzione di costruire una “macchina pensante” risaliva all’aprile del 1955, quando Ceccato e Maretti annunciarono la loro idea durante un convegno organizzato da Giovanni Emanuele Barié, fondatore e direttore dell’Istituto di Filosofia presso l’Università Statale di Milano, il quale commentò così l’intervento:

Ceccato aveva parlato della macchina cibernetica, capace di tradurre un latinetto di seconda media, o di radiotrasmettere una partita di calcio. Si trattava della fabbricazione dell’homunculus. [...] La fabbricazione dell’homunculus dovrebbe preludere alla fabbricazione di Dio. Se si prevede che l’uomo potrà andare sulla

²⁰⁵ Ceccato 1988.

²⁰⁶ A riprova dell’atteggiamento ambivalente di Sinisgalli, basti considerare che nei suoi primi quattro anni di vita *Civiltà delle macchine* ospitò oltre venti articoli su vari aspetti della cibernetica. Tra di essi, vi era *Costruzione di un homunculus*, in cui Giuseppe Vaccarino tracciava le linee guida per la costruzione di una macchina “in grado di servirsi delle nostre parole nelle situazioni e per le situazioni in cui noi ce ne serviamo”. Anziché essere semplicemente fonetica, la macchina avrebbe dovuto esprimere significati per mezzo di un processo semantico alle prese con simboli. Tale progetto era solo una delle possibili applicazioni delle analisi linguistiche condotte dalla Scuola operativa italiana, i cui articoli trovavano ampio spazio nel bimestrale di Sinisgalli. Cfr. Pogliano 2011: 355.

luna fra una trentina di anni, certo per la fabbricazione di Dio bisognerà stanziare qualche secolo. Ma è questione di tempo, come si vede; di misura.²⁰⁷

All'inizio del 1956, sempre su *Civiltà delle macchine*, apparve un articolo intitolato *La grammatica insegnata alle macchine. Note di Silvio Ceccato*, tratto dalle considerazioni esposte da Ceccato e Maretti al *Third London Symposium on Information Theory* intorno ai risultati da loro raggiunti nello studio della traduzione meccanica:

Sarebbe davvero miracoloso se, in una situazione in cui chi costruisce non può nemmeno contare su una definizione di 'linguaggio', saltasse fuori una macchina che addirittura sostituisce una lingua con un'altra, cioè traduce. No, la traduttrice meccanica non è stata fabbricata. [...] La decisione di costruire la traduttrice meccanica avrà dunque ripercussioni sia sulla tecnica e gli atteggiamenti ingegnereschi, sia sulla tecnica e gli atteggiamenti del filosofo.²⁰⁸

Adamo II era dunque un'occasione unica per realizzare il progetto di traduzione automatica avviato da Ceccato, basato sull'analisi dell'attività mentale in termini di operazioni. Un frammento della macchina, frutto del lavoro congiunto di Ceccato e Maretti, fu presentato l'8 aprile del 1956 nel settore del Gruppo Finmeccanica alla *Mostra internazionale dell'automatismo*, e il suo funzionamento esposto da Maretti nel numero 3 di *Civiltà delle macchine* dello stesso anno. Maretti esordiva illustrando brevemente il processo che aveva portato la Scuola Operativa Italiana a

²⁰⁷ Bontadini 1955: 301.

²⁰⁸ Ceccato 1956: 46.

concentrarsi sempre di più sugli studi cibernetici e linguistici. In particolare, era proprio il linguaggio a fornire la chiave d'accesso allo studio dell'attività cerebrale, in quanto suo prodotto isomorfico:

Svolgendo una delle sue analisi operative Ceccato aveva infatti trovato che quando noi si nomina una cosa, questo avviene in quanto vengono poste in rapporto due cose, non importa quali, rappresentandosi quella fatta per prima mentre si è attivi per la seconda. La semanticità, cioè, è data da un rapporto e non da un segnale.²⁰⁹

Assumere gli oggetti del mondo (esterni o interni) quali entità separate dal nostro essere comportava una loro collocazione spazio-temporale forzata, mentre alle attività “non spetta alcun posto e richiedono invece due momenti”. Concepire i prodotti della mente come attività da studiare nei termini di una funzione organica meccanizzabile avrebbe invece permesso di dare la giusta collocazione a concetti quali, ad esempio, quello di “generale” e di “particolare”, in quanto prodotti dell'attività del confrontare e non più intesi come entità fisiche estrinseche dall'osservatore. Visto il poco tempo a disposizione²¹⁰, venne realizzata solo una parte della macchina (il “frammento” di cui sopra) e, poiché *Adamo II* nasceva “tutto testa”, la scelta delle funzioni che avrebbe dovuto eseguire ricadde sulle 23 categorie mentali più frequenti e importanti, in un chiaro omaggio a Kant. Il “cervello elettronico” di Ceccato e Maretti era anche in grado di formulare

²⁰⁹ Maretti 1956: 26.

²¹⁰ La costruzione effettiva di *Adamo II* iniziò appena tre mesi prima l'inizio della mostra.

tre celebri frasi, il cui contenuto era puramente mentale: “Cogito ergo sum” di Cartesio, “Verum et factum convertuntur” di Vico e “Das Für-sich-sein im Anderssein ist der Prozess” di Hegel. L’intenzione era di mostrare l’attività correlatrice del pensiero, così come le sue designazioni con la correlazione linguistica, avvalendosi di un dispositivo elettronico per la riproduzione di questi processi:

Quali circuiti avremmo adoperato le lampadine, in modo da rendere visibile e seguibile immediatamente, attraverso il loro accendersi e spegnersi, il gioco temporale combinatorio. I comandi ai circuiti dovevano venire trasmessi da interruttori a nastro, mossi da motorini elettrici. E fu così che io trascrissi i risultati della analisi operativa, sia delle singole combinazioni mentali sia dei tre pensieri, risultati già designati con gli esse, in una scrittura atta a servire da interruttore a nastro: i trattini neri in corrispondenza del materiale conduttore, per chiudere i circuiti, ed i trattini bianchi in corrispondenza del materiale isolante, per aprire i circuiti.²¹¹

La presentazione di *Adamo II* suscitò un notevole interesse da parte dei *media*, concentrando l’attenzione generale su quello che per i tempi era un tema abbastanza inusuale.

Dino Buzzati intitolò il proprio articolo per il *Corriere della Sera* *Una macchina che pensa costruita da due italiani*, soffermandosi in particolare su un’osservazione fatta a Ceccato durante il congresso:

²¹¹ Maretti 1956: 31.

Se almeno in linea teorica si può costruire una macchina che riproduce il pensiero umano, se ne dovrebbe dedurre che l'uomo è governato da leggi meccaniche! E la sua libertà andrebbe a farsi benedire!²¹²

Forse il tono era un po' troppo sensazionalistico, ma inquadrava perfettamente una delle maggiori inquietudini generate da *Adamo II*: la fobia nei confronti di un meccanicismo esasperato, che sembrava negare all'essere umano il suo carattere di unicità nel mondo. Ceccato fece probabilmente del suo meglio per fugare qualunque dubbio a riguardo, affermando che “le macchine di questo genere aiuteranno l'uomo a comprendere se stesso e, non fosse per questo, a renderlo più responsabile delle sue azioni”.²¹³ Non una de-umanizzazione dell'uomo quindi, ma la costruzione di simulacri a nostra immagine che avrebbero permesso una maggior comprensione di quell'unicità che sembrava tanto minacciata. La creatura di Ceccato e Maretti non lasciò Buzzati indifferente, e l'idea che una macchina potesse riprodurre il pensiero umano (con tutte le implicazioni del caso) s'insinuò sempre più nella sua immaginazione. Prima a puntate su *Oggi*, e poi raccolto in volume da Mondadori, nel 1960 venne pubblicato *Il grande ritratto*, un romanzo di fantascienza che anticipava quelli che sarebbero diventati alcuni *topos* del genere.²¹⁴ L'influenza di

²¹² Ceccato 1956: 32.

²¹³ Ibid.

²¹⁴ Per quanto Buzzati non disponesse sicuramente della conoscenza “tecnica” tipica di altri scrittori di fantascienza, è curioso notare come il libro precedesse alcuni classici del genere come *Ma gli androidi sognano pecore elettriche?* (forse meglio conosciuto per il suo adattamento cinematografico, *Blade Runner*) di Philip K. Dick e *Il ciclo della fondazione* di Asimov, rispettivamente del 1966 e del 1968.

Ceccato è ulteriormente confermata dalla citazione delle “formule di Ceccatieff” presente nel libro. Ne *Il grande ritratto* il fisico Endriade costruiva, in una deserta valle di montagna, una colossale macchina battezzata “Numero Uno” fatta di cupole, casematte, cunicoli, sotterranei, antenne. Al centro di essa risiedeva un grosso uovo elettronico, l'anima di “Numero Uno”, che aveva il fascino e la perfidia di Laura, la moglie di Endriade morta in un incidente stradale.²¹⁵

Sempre sul *Corriere della Sera*, Ugo Maraldi esprimeva diffidenza sulla possibilità che le macchine potessero riprodurre, oltre che i fenomeni del corpo e della mente, anche quelli della coscienza umana. Nelle intenzioni dei suoi creatori *Adamo II* avrebbe dovuto infatti manifestare una forma, seppur abbastanza primitiva, di auto-consepevolezza, presupposto ritenuto necessario per qualunque attività di tipo mentale. Utilizzando come meccanismi principali una sorgente e un utilizzatore, la rappresentazione grafica delle loro interazioni avrebbe permesso la descrizione di qualsiasi stato mentale²¹⁶:

La soluzione al nostro problema apparve con l'analisi del “singolare”. Qui figuravano innegabilmente due pezzi (cioè due simboli) ben distinti fra loro; e si trattava di una combinazione tripla, in quanto uno di questi compariva due volte, sia prima che dopo l'altro. A questo punto non era difficile riconoscere anche che cosa fosse questo pezzo ripetuto: era lo stato di vigilanza, di attenzione non

²¹⁵ Alcuni di questi temi saranno ripresi in un altro libro di Buzzati, *Un amore*.

²¹⁶ Prese ad esempio la coppie aA-bA e aA-aB (con le lettere minuscole a indicare le sorgenti e le maiuscole gli utilizzatori), esse avrebbero rappresentato i concetti di “stesso” e “altro”.

focalizzata, ciò che di solito viene chiamato “coscienza”, in espressioni come “perdere in coscienza”, “acquistare la coscienza”. E l’altro pezzo? Ceccato mi telefonò una mattina trionfante. L’altro pezzo era il “qualcosa”, e si poteva ottenere per sovrapposizione di due stati di coscienza.²¹⁷

Le altre reazioni all’esperimento (passate in rassegna da Ceccato in appendice all’articolo di Maretti) oscillavano dal più assoluto entusiasmo fino a una certa pessimistica inquietudine per il futuro.²¹⁸ Se l’intenzione era far parlare di sé, per puro spirito di provocazione o al fine di ottenere finanziamenti per le proprie ricerche (o per entrambe le cose), l’obiettivo era stato sicuramente raggiunto.²¹⁹

Adamo II rientrava in quel campo di ricerca che lo stesso Ceccato avrebbe chiamato, anni dopo e con maggiore consapevolezza, “terza cibernetica”. Ceccato coniò il termine nei primi anni Sessanta proprio per distinguere i suoi studi da quelli wieneriani, interessati alla costruzione di macchine che simulavano una prestazione umana (ma non le catene di operazioni o i

²¹⁷ Maretti 1956: 27.

²¹⁸ Ceccato 1956: 32.

²¹⁹ Esiste un aneddoto che permette di inquadrare meglio la personalità di Ceccato. Scriveva di lui Pierluigi Ridolfi: “Un giorno lo trovai nel suo studio di via Festa del Perdono, a Milano, sommerso da tante foglie di lattuga: mi disse che stava studiando la lattughità, cioè la proprietà intrinseca della forma della lattuga rispetto a quella del cetriolo, e che stava progettando una macchina in grado di riconoscere qualunque tipo di ortaggio. Questa consisteva in una tramoggia nella quale gettare il vegetale, delle cellule fotoelettriche, un computer, un programma. La risposta sarebbe stata: “è lattuga” oppure “è cetriolo”. Gli chiesi perché due vegetali e non piuttosto una sfera e un cubo, sarebbe stato più facile. Ma Ceccato mi fulminò con la sua risposta: “Una sfera non fa notizia, ma vedrai la lattuga!!!” E, infatti, qualche settimana dopo, il *Giorno di Milano* titolò a 9 colonne: “Uno scienziato italiano costruisce una macchina intelligente che riconosce le verdure”. Ceccato sapeva perfettamente che si trattava di stratagemmi per far parlare di sé; sta di fatto che anche in questo modo per anni riuscì a ottenere dei fondi per sopravvivere (perfino dalla Nasa per tradurre automaticamente dal Russo!).” Cfr. Ridolfi 2007: 24-25.

processi umani che la producono), sia dalla bionica, interessata anch'essa ai processi, ma esclusivamente a quelli neurali e biologici.²²⁰ La “cibernetica della mente” (o logonica) di Ceccato si concentrava quindi su quell'analisi delle operazioni mentali trascurata dalla “prima” cibernetica (considerata “comportamentistica”) e irraggiungibile per la “seconda”.²²¹ Lo stesso *Adamo II* presupponeva l'analisi dei meccanismi che danno vita al pensiero, scomposti in passaggi discreti e poi riprodotti dalla macchina.²²² Non a caso Maretti, durante la presentazione dell'automa, prendeva le distanze da Shannon, Ashby e Walter, che avevano cercato di “costruire meccanismi in grado di dare certe risposte tropistiche a certe stimolazioni”, ma senza mai porsi il problema della correlazione tra i prodotti della mente (come, ad esempio, il linguaggio) e il cervello: *Adamo II* era per i suoi autori “la prima macchina pensante e parlante che sia stata prospettata”.²²³

Sarebbe probabilmente un po' avventato vedere in *Adamo II* il primo esempio di AI sviluppato in Italia, poiché la macchina nel suo complesso era quanto di più lontano dall'AI classica ci si possa immaginare. E' più corretto dire che l'interesse di Ceccato verso la cibernetica si estrinsecò in un ideale quasi leibniziano rivolto a formulare una metodologia che permettesse di rendere sempre meno ambiguo e più efficace il linguaggio: lo stesso *Adamo II* nasceva come semplice esemplificazione a scopo

²²⁰ Si veda per esempio l'articolo del 1964 *Contra Wiener, pro Wiener*.

²²¹ Ceccato 1972.

²²² Bozzo 1996: 111-112.

²²³ Maretti 1956: 27.

dimostrativo delle possibilità offerte dall'analisi operativa delle attività mentali e linguistiche.²²⁴

Ma in retrospettiva e nel quadro della sua visione modellistica, che vedeva nei dispositivi elettronici un modello per lo studio della mente umana, era proprio l'attività di Ceccato come linguista (che nella costruzione di *Adamo II* giocò un ruolo fondamentale, pur concretizzandosi in una macchina tipicamente "cibernetica"), quella che con il tempo si sarebbe rivelata la più innovatrice.

Abbandonata la costruzione del "cronista meccanico", che avrebbe dovuto essere la diretta evoluzione di *Adamo II*, gli anni Sessanta videro Ceccato dedicarsi quasi esclusivamente allo studio delle difficoltà insite nell'approccio tradizionale alla traduzione automatica. Fece grande scalpore l'*Almanacco Bompiani* del 1962, dedicato alle applicazioni dei calcolatori elettronici alle scienze morali e alla letteratura, al quale Ceccato contribuì con l'articolo *La storia di un modello meccanico dell'uomo che traduce*.

Scrivendo Ceccato:

Si parla di traduzione meccanica per alludere ad un complesso di ricerche che hanno quale primo oggetto lo studio dell'uomo che traduce e solo quale secondo oggetto la costruzione di una macchina che traduce, una macchina che assolve quindi funzioni modellistiche, cioè riproduce nei limiti del possibile il lavoro di un uomo traduttore. E si sa bene che cosa voglia dire studiare un uomo che traduce. Vuol dire importare sì una indagine di tipo linguistico, ma ben più estesa

²²⁴ Somenzi 1993: 181.

ed approfondita di quanto non si intenda di solito, con connessioni in campo psicologico ed infine filosofico, che permettono di considerare il linguaggio non soltanto come un prodotto già acquisito, ma anche in tutto il dinamismo che accompagna lo svolgersi del pensiero.²²⁵

Ceccato sarebbe stato successivamente riconosciuto come uno dei pionieri dei moderni approcci all'analisi del linguaggio naturale e alla traduzione automatica: studi che a loro volta sarebbero stati assimilati al più ampio campo di ricerca dell'AI. Nello stesso periodo in cui egli pubblicava il suo articolo per l'almanacco, nelle ricerche anglosassoni sull'elaborazione del linguaggio naturale si era ormai abbandonato da tempo l'obiettivo della cosiddetta "traduzione completamente automatica di alta qualità" (*fully automatic high quality machine translation*), e si cercavano nuove strade per affrontare la gestione del linguaggio naturale con i calcolatori. Il motivo di questo fallimento era stato discusso da un altro pioniere del settore, Yehoshua Bar-Hillel, durante la prima *International Conference on Machine Translation* tenuta al MIT nel 1951.²²⁶ Data ad esempio la frase "il cane si è inceppato", il parlante di lingua italiana sa che qui con "cane" non ci si riferisce a un animale (CANE1), ma al percussore di un'arma da fuoco (CANE2). Da qui le difficoltà per una macchina di tradurre correttamente la frase in inglese, dove CANE1 è *dog* mentre CANE2 è *cock*. Casi del genere potevano moltiplicarsi a piacere, a conferma del fatto che una buona

²²⁵ Ceccato 1962: 122-123.

²²⁶ Hutchins 1986: 122

traduzione interlingua, ma in generale una buona comprensione delle lingue, non poteva prescindere dai significati suggeriti dal contesto e dalla conoscenza implicita nel lessico dei parlanti.²²⁷ Secondo lo stesso Bar-Hillel, era impossibile riuscire a rappresentare queste caratteristiche in un programma per calcolatore, e sarebbe stato preferibile un obiettivo forse meno ambizioso, ma più praticabile, che non poteva comunque prescindere dall'intervento dell'uomo:

A mixed MT (machine translation) [...] in which a human brain intervenes [...] either at the beginning of the translation process or the end, perhaps at both, but preferably not somewhere in the midst of it.²²⁸

Gli studi di Ceccato di quegli anni si concentravano proprio sui problemi del ragionamento di senso comune e del ruolo del contesto nell'interpretazione delle frasi e della loro disambiguazione, nonché del carattere implicito dell'informazione usata dai parlanti. In questo senso, le sue “sfere nozionali” erano un tentativo di affrontare l'insieme di queste difficoltà, spesso insormontabili per le macchine, ma relativamente facili da superare per gli esseri umani. Le sfere nozionali erano reti di concetti collegati da nessi relazionali a vario livello di generalità o specificità, come tutto-parte, genere-specie, classe-istanza, o da nessi relazionali come quelli temporali o spaziali. Esse anticipavano di fatto i successivi sistemi di rappresentazione della conoscenza sviluppati in AI a partire dalla seconda

²²⁷ Cordeschi 1996: 167

²²⁸ Bar-Hillel 1951: 236.

metà degli anni Sessanta, come ad esempio le reti semantiche di Ross Quillan, e successivamente i *frames* di Minsky e gli *scripts* di Roger Schank.²²⁹

Che la “terza cibernetica”, pur sviluppandosi da premesse originali, non fosse del tutto estranea all’AI degli inizi lo confermava lo stesso Ceccato, che a partire dal 1965 ne curò la voce omonima per *l’Enciclopedia della Scienze e della Tecnica* edita dalla Mondadori:

Sotto il nome di intelligenza artificiale si comprende l’oggetto di un gruppo di ricerche che, prendendo in considerazione le attività umane chiamate intelligenti, o mentali, o superiori, si propone la costruzione di meccanismi che ripetano queste attività.²³⁰

Sebbene Ceccato facesse rientrare sotto questa definizione la costruzione di reti neurali (“si sa che la nostra attività mentale ha quale organo il sistema nervoso”), essa comprendeva anche gli studi linguistici e sulla traduzione automatica cui ormai si stava dedicando da oltre un decennio:

Si sa che il linguaggio è legato all’attività mentale, al pensiero; per cui, se la macchina è capace adoperare come noi le espressioni linguistiche, se ne deduce che anch’essa svolge attività mentale, di pensiero, e quindi è intelligente.²³¹

Tutto ciò sarebbe stato possibile ricorrendo a un calcolatore dotato di elaborazione binaria, dal momento che qualunque operazione elementare

²²⁹ Cordeschi 2002: 193.

²³⁰ Ceccato 1965: 282.

²³¹ Ibid.

poteva essere ricondotta al funzionamento di un organo a due stadi.

La descrizione di Ceccato non era immediatamente riconducibile all'AI classica, poiché comprendeva metodologie che, già 9 anni prima, i pionieri dell'AI avevano giudicato inconcludenti. Non a caso, nella bibliografia dell'articolo, accanto ad alcuni dei più importanti scritti di Minsky e Simon, venivano citati anche anche gli *Automata Studies* e *A Logical Calculus of the Ideas Immanent in Nervous Activity* di McCulloch e Pitts. Di contro, la differenza tra i due approcci era illustrata con una distinzione piuttosto netta: mentre le reti neurali si limitavano a meccanizzare un comportamento arbitrariamente fissato, altri studi, come quelli sulle manifestazioni linguistiche e sulla traduzione automatica, necessitavano di un'analisi psicologica, della costruzione di un modello mentale che teneva conto delle sole funzioni superiori. Ma per Ceccato era proprio questo il problema: l'assenza nella macchina di una qualsiasi attività mentale e rappresentazionale. Da qui l'illusione che una macchina "pensi", nonostante il suo operare interamente fisico.

Concludeva Ceccato:

Si tratta di ricerche che danno buone speranze, anche se alla meccanizzazione delle nostre attività di osservazione, di categorizzazione e soprattutto di pensiero e di linguaggio si oppone per la mancanza di un principio per la costruzione di un aspetto essenziale alle manifestazioni dell'intelligenza, quello della memoria: una

memoria sia non già a magazzino, ma con le caratteristiche di quella umana, che è selettiva, associativa e soprattutto propulsiva.²³²

Oltre a problemi di carattere epistemologico come la creazione di un mondo rappresentazionale di conoscenza, Ceccato aveva individuato un'altra questione abbastanza delicata: la realizzazione di una memoria artificiale "elastica", capace di simulare le principali caratteristiche di quella umana. Studi a riguardo sarebbero stati condotti, molti anni dopo, dal biologo Steven Rose, che rifiutava un modello tratto dall'elaborazione dell'informazione a favore di una comprensione più dinamica della memoria come processo biologico, caratterizzato da crescita strutturale, riconfigurazione ambientale e decadenza senile.²³³ Diversamente dal computer, che ripete il procedimento algoritmico con precisione e indipendentemente dall'orologio biologico, Rose sosteneva che la memoria umana è ridondante, ma che probabilmente sono queste caratteristiche a permetterle di esercitare determinate funzioni (come l'esprimere un giudizio) con facilità estrema.²³⁴

4.2 1962

Nel dicembre del 1961 la casa editrice Bompiani pubblicava il proprio almanacco letterario annuale, la cui copertina, consistente di un primo piano della memoria a nuclei ferritici del calcolatore Elea 9000, esaltava un titolo

²³² Ivi.: 283.

²³³ Rose 1993: 117

²³⁴ Agazzi, Fortunati 2007: 490

abbastanza insolito: “Le applicazioni dei calcolatori elettronici alle scienze morali e alla letteratura”.

Non solo, quindi, un repertorio dei fatti più rilevanti avvenuti nel corso di quell’anno in campo letterario, come ci si sarebbe potuto aspettare da una pubblicazione di questo tipo, ma anche un’introduzione a un tema (l’elettronica) conosciuto solo da una ristretta cerchia di specialisti e percepito come lontanissimo dalla letteratura e dalle scienze morali.

L’introduzione (presumibilmente scritta dal curatore Sergio Morando) sottolineava come tutto ciò facesse parte dello spirito dell’almanacco, che giusto l’anno prima aveva pubblicato un’antologia delle letterature meno tradotte:

L’almanacco cerca di dare al lettore quegli aspetti inediti, a volte curiosi, a volte inquietanti, che sembrano annunciare qualche nuova direzione e dimensione del futuro.²³⁵

Seguiva l’articolo di Rinaldo De Benedetti su *Il calcolo proposizionale ovvero l’algebra delle idee*, che tracciava una breve storia degli sviluppi della logica simbolica da Euclide a Bertrand Russell (passando per Leibniz e George Boole), spiegando così la base logico-matematica della nuova disciplina. L’articolo analizzava l’utilità e l’interesse del calcolo proposizionale attuato mediante la logica simbolica, che poteva trovare utilizzo non solo nella risoluzione di problemi a carattere aziendale e

²³⁵ Aa. Vv. 1962: 7.

organizzativo, ma anche in quella di “problemi logici complicati, veri rompicapo, che non sempre ma sovente si presentano alla pratica e si riducono a proporzioni maneggevoli”.²³⁶ La logica simbolica trovava insomma “la sua applicazione più importante e corposa nella costruzione delle macchine calcolatrici, dei cervelli elettronici”.²³⁷

La successiva esposizione trattava dei fondamenti dell'algebra binaria e prendeva le mosse dall'analisi della maieutica socratica presente nel dialogo platonico *Il sofista*, in cui il metodo dialogico utilizzato da Socrate nei confronti dell'interlocutore veniva ricondotto a una strutturazione di domande che presupponevano risposte in termini binari:

Ebbene, il cervello elettronico ragiona un poco come gli interlocutori del *Sofista*: esso sceglie solo tra due possibilità: *si o no*, 1 oppure 0, *vero* oppure *falso*, salvo che compie queste operazioni a velocità vertiginosa e per un numero altissimo di scelte successive.²³⁸

Michele Pacifico scriveva poi de *I nuovi Gutenberg. Linguistica ed elettronica nel mondo, oggi*, il tutto sotto l'insegna di una rassicurante citazione di Wittgenstein, che aveva probabilmente lo scopo (un po' come tutta la prima parte dell'almanacco) di esorcizzare le paure degli umanisti: “Chiedersi se una macchina possa pensare è un po' come chiedersi che colore abbia il numero tre”. Pacifico si soffermava sul problema della

²³⁶ De Benedetti 1962: 91.

²³⁷ Ibid.

²³⁸ Aa. Vv. 1962: 92.

codifica, cioè della simbolizzazione dell'informazione, affermando che in linea di principio nulla impediva che tale processo potesse avere come destinatario un calcolatore. Il tutto, ovviamente, presupponendo un'analisi comportamentistica della macchina: dal momento che essa era provvista di memoria e agiva seguendo le regole della logica, allora era possibile, almeno in linea di principio, "l'idea di una possibile collaborazione tra linguistica ed elaboratore elettronico".²³⁹ Venivano poi passati in rassegna i centri che si dedicavano alle "ricerche linguistiche, intese nel senso più vasto e comprensivo di questo termine".²⁴⁰

Non mancarono neanche i resoconti di esperimenti già compiuti (o in corso d'opera) sull'applicazione dei calcolatori ai domini più vari, cui l'almanacco dedicò tutta la parte centrale.

Il gesuita padre Roberto Busa, ad esempio, illustrava le applicazioni del metodo statistico ai fenomeni linguistici e all'analisi lessicale. Busa non era nuovo a questo tipo di ricerche: risaliva al 1946 l'idea di utilizzare i calcolatori per la produzione di *indices* sistematici, e già a partire dal 1949, con le prime macchine elettromeccaniche a schede dell'IBM, iniziava la colossale impresa dell'*Index Thomisticus*, comprensivo di tutte le opere di S. Tommaso d'Aquino (in 56 volumi).²⁴¹

Il direttore della Centrale Calcoli Elettronici della Fiat Stanislao Valsesia esponeva le linee guida per la costituzione di un enorme *corpus* di dati, una

²³⁹ Pacifico 1962: 101.

²⁴⁰ Ibid.

²⁴¹ Morenti 2001: 11.

futura “biblioteca elettronica” in cui sarebbe stato possibile trovare tutti gli articoli desiderati a partire da certe parole chiave che li designavano (praticamente un proto-motore di ricerca). Il professor Carlo Tagliavini descriveva *L'automazione nelle ricerche fonetiche (ma anche morfologiche e sintattiche)* condotte sulla lingua contemporanea nell'Istituto di Glottologia dell'Università di Padova, mentre Silvio Ceccato, come già detto, scrisse *La storia di un modello meccanico dell'uomo che traduce*, citando la realizzazione da parte dell'IBM di una macchina traduttrice per il governo americano.

Circa gli effetti di questa nuova dialettica, l'almanacco ospitava un'inchiesta sulle cosiddette “due culture” (scientifica e umanistica), consistente in quattro domande rivolte a quattordici filologi e letterati:

1. Come ella saprà, presso alcuni istituti universitari europei sono in corso da diversi anni ricerche filologiche e linguistiche nelle quali si ricorre all'impiego dei moderni strumenti di elaborazione elettronica delle informazioni, i cosiddetti “cervelli elettronici”. [...] Ritiene che l'intervento della macchina nel campo della filologia sia un fatto “positivo” o no?

2. A giudizio di alcuni, l'utilizzazione delle macchine elettroniche nelle ricerche lessicologiche e letterarie può portare, anzi porterebbe inevitabilmente, a un mutamento profondo nei metodi e anche nei principi ispiratori di tali ricerche.

Altri invece ritiene che dal ricorso ai “cervelli elettronici” non ci si debba aspettare altro che uno sveltimento e un alleggerimento delle procedure classiche [...]. Quale di questi due punti di vista ritiene di poter condividere?

3. In particolare quale contributo, secondo Lei, le nuove tecniche per l'elaborazione dei dati in campo lessicologico possono dare alla critica stilistica?
4. La collaborazione in atto [...] tra il mondo della tecnica e quello delle scienze umanistiche può, a Suo parere, svilupparsi fino a superare l'attuale frattura tra scienza e umanistica (fra le "due culture" per dirlo con Snow) oppure non è altro che uno dei sintomi, sempre più frequenti, dell'assorbimento delle scienze morali nelle tecniche della civiltà meccanizzata contemporanea?²⁴²

Le risposte degli interessati erano generalmente concordi nel valutare positivamente l'utilizzo delle tecniche di elaborazione elettronica negli studi di carattere filologico, ma anche nel considerare strumentale e del tutto circoscritto a una questione puramente "tecnica" il contributo che le metodologie in questione potevano apportare alla critica stilistica. Anche riguardo alla seconda domanda, le risposte tendevano per lo più a escludere un possibile mutamento dei metodi e dei principi ispiratori delle ricerche letterarie dovute alle nuove tecnologie. L'unica voce fuori dal coro fu Gianfranco Contini, secondo cui il "significato euristico" della macchina, derivato dal suo poter condurre indagini quantitative fino a quel momento negate, si sarebbe alla fine rivelato determinante: gestione e metodo avrebbero finito con l'influenzarsi a vicenda.²⁴³

Dato il contesto, Contini sembrava attribuire al termine "euristico" lo stesso significato che esso assume per alcuni teorici dell'AI: "Un'euristica è una

²⁴² Ibid.: 143.

²⁴³ Ibid.: 144.

tecnica che migliora l'efficienza di un processo di ricerca".²⁴⁴

Dopo l'inchiesta, l'almanacco riportava il resoconto dell'esperimento "poetico" compiuto da Nanni Balestrini utilizzando un calcolatore IBM, su cui ci si soffermerà in seguito.

L'articolo successivo, firmato da Franco Lucentini, portava il titolo (piuttosto altisonante) di *Automatopoietica*. Con esso, l'almanacco passava da un discorso fino a quel momento incentrato sulle applicazioni degli elaboratori in linguistica a uno più di carattere letterario e umanistico, tracciando una sorta di storia letteraria del mito dell'"uomo artificiale".

L'esposizione di Lucentini tentava di illustrare la genesi dell'ambiguità di fondo insita negli automi, e del rapporto spesso conflittuale che l'uomo instaura con le macchine da lui stesso create.

L'analisi degli automi presenti nella letteratura greca e latina (ad esempio nell'*Iliade*) spinse Lucentini ad affermare che negli scritti dell'età classica automi e androidi erano presenti solo in contesti perfettamente tranquilli e rispettabili, utilizzati più che altro come strumenti degli dei.²⁴⁵

Ciò che mancava non era la condanna morale verso l'uomo-fantoccio (mosso da altro), ma quella di tipo metafisico e religioso verso l'uomo-automa, che si muove da sé. Anche nel *De anima* di Aristotele l'anima ispirava e guidava i movimenti del corpo, per non lasciare che quest'ultimo fosse sopraffatto del caso, ma non vi era ancora nessuna condanna della

²⁴⁴ Rich 1986: 51.

²⁴⁵ Lucentini 1962: 154.

materia in quanto tale. Secondo Lucentini, ciò sarebbe avvenuto solo dopo, con l'incontro tra la tradizione greco-latina e quella giudaico-cristiana, quando la materia cessò di essere passività necessitante lo spirito per animarsi e divenne negazione attiva e violenta di quello stesso spirito. Dopo aver attraversato rapidamente il Medioevo, Lucentini citava lo straordinario sviluppo di automi e androidi (per lo più innocui e puramente "estetici") nelle corti europee del Cinquecento e del Seicento; sviluppo che sarebbe andato di pari passo con il declino del concetto di anima nelle filosofie, ad esempio, di Cartesio e Hobbes:

Man mano che i filosofici automi avanzano, aumenta l'apprensione del pubblico: che è ormai troppo assuefatto al possesso dell'anima, per contemprarne impavido la rapida recessione.²⁴⁶

E proprio in questo clima d'isteria quasi sottocutanea si sarebbe assistito alla prima ribellione furiosa di un automa nei confronti della razza che lo aveva generato: il Golem di Rabbi Loew.²⁴⁷ A partire dall'Ottocento, dunque, la paura degli automi iniziò a dilagare, e non servì a nulla che, per dissipare il panico, razionalisti come D' Alembert dedicassero cinque intere colonne del primo volume dell' *Encyclopedie* a descrivere in termini rassicuranti uno degli inoffensivi congegni di Vaucanson. Al contempo,

²⁴⁶ Ibid.: 157.

²⁴⁷ A fine articolo veniva citato un altro esempio di automa ribelle, ovvero *L'apprendista stregone* di Goethe. Gli stessi esempi vennero utilizzati anche da Wiener nell'articolo *Some Moral and Technical Consequences of Automation* del 1960, dove si sottolineava come una comprensione critica delle macchine dovesse sempre procedere di pari passi con le loro effettive capacità, al fine di non sottovalutarle ed evitare così conseguenze disastrose.

però, s'iniziava a delineare anche la figura dell'automa "romantico": non un uomo-macchina, ma un *Massimo Uomo*, l'automa spirituale di Swedenborg che sgominava i mostri del materialismo e avrebbe portato al *Frankenstein* di Mary Shelley.

Torniamo all'articolo di Balestrini. La presentazione dell'articolo recitava:

Letteratura e arte hanno nell'ultimo cinquantennio costantemente prestato una attenzione vivissima ai fondamenti dei propri processi immaginativi e costruttivi, individuabili e riassumibili nelle successive fasi di decomposizione dei materiali precostituiti, e di ricomposizione in un risultato creativo.²⁴⁸

Balestrini elencava le tappe principali della storia dell'arte combinatoria, citando (tra gli altri) Mallarmé, Ungaretti e Joyce, "legittimando" e riconducendo il proprio esperimento all'interno di una tradizione storica prestigiosa.

Ma come giustificare l'utilizzo di un calcolatore, di una "macchina", per la creazione di qualcosa che veniva (e viene tuttora) considerata prerogativa dell'ingegno e della sensibilità umana, la poesia? Era lo stesso Balestrini a rispondere a questa domanda:

La utilità e la legittimità dell'impiego dei metodi e dei mezzi messi a disposizione dalla scienza e dalla tecnologia più progredita, intendendoli come integrazione

²⁴⁸ Balestrini 1962: 145.

dell'opera di creazione letteraria e artistica, si manifestano in accordo al nostro appartenere ad una civiltà industriale.²⁴⁹

Il suo scopo dichiarato non era quello di imitare i procedimenti umani, come le altre “prove sul linguaggio svolte nell'ambito della cibernetica” illustrate nell'almanacco, ma sfruttare le capacità del calcolatore per risolvere con estrema rapidità alcune complesse operazioni inerenti alla tecnica poetica.²⁵⁰ Poeti e umanisti, insomma, non dovevano sentirsi minacciati, dal momento che la creatività umana non ne risultava in nessun modo svilita.

L'esperimento elettronico di Balestrini traeva la sua origine dall'installazione, avvenuta nel 1960, di un elaboratore 7070 dell'IBM presso la sede centrale della Cassa di Risparmio delle Province Lombarde. Alberto Nobis, tecnico dell'IBM, era il responsabile di quell'installazione e autore materiale del programma utilizzato nell'esperimento. Nonostante l'accesso al calcolatore fosse probabilmente precluso al personale non autorizzato, i tecnici dell'IBM (così come quelli dell'Olivetti) si dimostrarono estremamente disponibili nei confronti degli umanisti che curarono l'almanacco (cosa che veniva ribadita nell'introduzione). Il resoconto di Balestrini, corredato da tabulati e diagrammi di flusso, era abbastanza esauriente per quanto riguardava il lato “tecnico” dell'esperimento:

²⁴⁹ Ibid.

²⁵⁰ Ibid.

E' stato predisposto un testo formato da tre brani tematicamente differenti, suddiviso in sintagmi (elementi) formati ciascuno da 2 o 3 unità metriche. Ciascun elemento è stato contraddistinto da un codice di testa e da un codice di coda, indicanti le possibilità sintattiche di legame tra due elementi successivi.²⁵¹

L'intervento del calcolatore consisteva nella composizione di una poesia di 6 strofe, ognuna formata da una diversa combinazione dei testi dati. Le strofe dovevano risultare di 6 versi ciascuna e ogni verso costituito da 4 unità metriche. I frammenti scelti da cui estrarre i sintagmi da combinare erano tratti dal *Diario di Hiroshima* di Michihito Hachiya, *Il mistero dell'ascensore* di Paul Goodwin e il *Tao te king* di Lao Tse. Dopo aver scomposto in unità minimali i materiali linguistici, essi venivano rimontati in un collage metodicamente preordinato che seguiva delle regole prestabilite, secondo le quali a ogni elemento doveva seguirne un altro che presentasse nel codice di testa una cifra corrispondente a una delle due formanti il codice di coda del primo. Il fatto che le combinazioni prodotte dal programma fossero il risultato di una formalizzazione preventiva degli input da elaborare (compiuta dal programmatore) confermava le parole di Balestrini a inizio articolo: la macchina non effettuava nessuna analisi semantica, ma si limitava a usare delle procedure che qualunque essere umano avrebbe potuto impiegare. Così, attribuiti a ogni sintagma i relativi "codici di testa" e "codici di coda", era già possibile prevedere in linea di massima quali sarebbero state le possibili combinazioni realizzabili,

²⁵¹ Ibid.

evitando l'accostamento di sintagmi troppo simili. Un'altra regola prevedeva l'impossibilità di sovrapporre sintagmi provenienti dallo stesso gruppo, garantendo così all'opera un certo grado di originalità e omogeneità. L'esperimento di Balestrini non tardò a suscitare delle reazioni, e la prima di esse si trovava all'interno dell'almanacco stesso. Alla parte monografica del volume contribuì anche Umberto Eco, che nel suo articolo *La forma del disordine* scrisse:

Questo è molto importante: in altra parte dell'Almanacco troveremo le poesie elettroniche di Nanni Balestrini. Con la complicità di un poeta e di un ingegnere programmatore, il cervello IBM ha sparato più di tremila variazioni dello stesso gruppo di versi, tentando tutte le combinazioni che le regole di partenza gli davano come possibili. Se andiamo a cercare tra i tremila risultati ne troveremo alcuni insulsi, altri (pochi, mi pare) di altissima temperatura lirica, che non avremmo esitato ad attribuire ad un cervello umano. Ma è proprio qui l'errore: questi pochi risultati elettissimi, probabilmente Balestrini sarebbe stato in grado di ottenerli da solo a tavolino; scelti i versi, poco ci voleva a metterli insieme nel modo più acconcio "alla maniera di", o comunque in armonia con certe correnti di gusto. L'opera del cervello elettronico, e la sua validità (se non altro sperimentale e provocatoria) consiste invece proprio nel fatto che le poesie sono tremila e bisogna leggerle tutte insieme. L'opera intera sta nelle sue variazioni, anzi nella sua variabilità. Il cervello elettronico ha fatto un tentativo di "opera aperta".²⁵²

²⁵² Eco 1962: 176.

Per quanto volutamente provocatorio, il commento di Eco non si concentrava tanto sull'analisi delle poesie prodotte da Balestrini, ma sul mezzo che egli aveva utilizzato per raggiungere il suo obiettivo e sulle sue implicazioni filosofiche (cosa che lo stesso Balestrini si era premurato di negare appena qualche pagina prima). Non solo il "cervello elettronico" poteva sostituirsi al poeta nella produzione artistica, ma la sua poesia, fatta di molteplici varianti simultanea, poteva fornire all'uomo contemporaneo "traduzioni immaginative" atte a comprendere e interpretare la realtà "scientifica" che lo circondava. Il calcolatore, pur con risultati nel complesso ben lungi dall'essere eccelsi e con tutte le dovute restrizioni, aveva provato a creare un'opera "aperta", diventando mezzo per interpretare la contemporaneità, e sarebbe stata questa la linea interpretativa delle recensioni all'esperimento di Balestrini.

Già sulle pagine dell'*Espresso* del 10 dicembre 1961 era possibile trovare un articolo intitolato *Bompiani ordina poesie a macchina*, con un occhietto che affermava "Il cervello elettronico entra nella storia della letteratura".²⁵³ Dello stesso tono fu anche l'articolo di Carlo Bo pubblicato sull'*Europeo* del 14 gennaio 1962: "Gli ingegneri del verso. Un calcolatore elettronico al posto dell'intelligenza del poeta". Bo recensiva i *Cent mille milliards de poèmes* di Raymond Queneau e l'esperimento elettronico di Balestrini, utilizzando entrambi gli avvenimenti letterari per un'analisi sul ruolo degli intellettuali nella società moderna:

²⁵³ Zanetti 1961.

E' un po' il quadro della nuova vita intellettuale: venti, venticinque anni fa il letterato conservava ancora un'aria romantica, rispondeva, bene o male che fosse, a una categoria ben definita nel periodo fra le due guerre, insomma era ancora un letterato puro. Oggi quel letterato non esiste quasi più, gli stessi uomini che si sono formati allora o addirittura avevano una posizione di primo piano, erano delle guide, vivono ritirati, difendono una ragione, un modo di vivere che non è più sentito. Il gusto sperimentale ha sostituito tutto il resto, e, di solito, non si fanno più calcoli a lunga scadenza: il letterato dei nostri giorni crede meno all'eternità della sua opera e ha accettato di dividere dentro di sé, nella parte più gelosa del suo cuore, una convinzione del suo tempo: cioè, tutto è votato alla luce di un giorno, basta provare, basta tentare, e quello che conta è essere presente intorno al bancone degli esperimenti.²⁵⁴

Gran parte delle idee esposte nell'almanacco sarebbero state sviluppate nel corso degli anni, ad esempio, da Giuseppe Gigliozzi, pioniere dell'Informatica umanistica in Italia che si dedicò, tra le altre cose, anche alla composizione automatica di poesie.²⁵⁵ Gli studi critico-letterari condotti tramite metodologie informatiche spinsero Gigliozzi (scomparso nel 2001) a esplorare le potenzialità dell'analisi automatica del racconto mediante tecnologie mutate dall'AI; negli anni Ottanta e Novanta del secolo scorso sperimentò e realizzò diversi modelli computazionali dei fenomeni narrativi, come la creazione di *story grammar* per fiabe e novelle e la descrizione

²⁵⁴ Bo 1962.

²⁵⁵ Per una rassegna di quei primi tentativi della *Computer Poetry*: D'Ambrosio 1986 (e la bibliografia sul dibattito internazionale ivi contenuta).

formale dei personaggi e dei ruoli narrativi, realizzando anche diversi programmi basati sul linguaggio LISP, come SEB e SEBNET.²⁵⁶ Esplorò inoltre le potenzialità dell'analisi quantitativa applicata al lavoro critico di un testo; si dedicò al problema teorico e pratico della codifica digitale dei testi e a quello degli studi di archivistica e biblioteconomia digitale, dove fu pioniere nella riflessione e nella sperimentazione sulle biblioteche digitali.²⁵⁷

4.3 1967

Nel 1965 comparve presso l'editore Boringhieri l'antologia *La filosofia degli automi*, curata da Vittorio Somenzi, che comprendeva la traduzione italiana degli articoli di Shannon, Turing e Wiener già apparsi in precedenza sulla rivista *Methodos*, oltre che alcuni testi fino a quel momento inediti come *The General and Logical Theory of Automata* di von Neumann e "Il libro delle macchine" tratto dal romanzo di Samuel Butler *Erewhon*.

Questa mescolanza fatta di antefatti storici e teoretici della cibernetica, oltre che di speculazioni sul futuro, catturò l'interesse di scrittori quali, oltre il già citato Buzzati, Italo Calvino, Primo Levi e Alberto Moravia, ma anche di giornalisti lontani da quel campo d'interessi, come Giorgio Bocca, i quali vi videro il segno di rottura della tradizione umanistico-idealistica italiana e

²⁵⁶ Ciotti, Crupi 2011: 269.

²⁵⁷ Ivi.: 9.

di un'apertura verso il metodo tipico della scienza anglosassone di affrontare i grandi problemi della filosofia, in particolare quello dei rapporti tra mente e corpo.²⁵⁸ Probabilmente fu questo il motivo principale per cui il XXI Congresso della Società Filosofica Italiana, tenuto nel 1967 presso l'Università di Pisa, fu dedicato al tema *L'uomo e la macchina*. L'idea venne ad Augusto Guzzo e ad altri filosofi italiani, tra cui Guido Calogero e Franco Lombardi, proprio per via delle reazioni che l'antologia di Somenzi stava suscitando non solo presso i *media*, ma anche negli ambienti umanistici.²⁵⁹ Tuttavia, altri fattori erano in gioco: la scelta di Pisa fu tutt'altro che casuale e rifletteva un sempre più sentito interesse intorno alla questione dei rapporti fra scienza, tecnologia, filosofia, economia, società e politica, particolarmente avvertita nel "biennio rosso" della contestazione e in piena guerra fredda.²⁶⁰ Proprio a Pisa, tra il 1954 e il 1960, era stata progettata e costruita la CEP (Calcolatrice Elettronica Pisana), uno dei primi calcolatori digitali attivi in Italia, suggerita dallo stesso Enrico Fermi.²⁶¹ Nello stesso periodo si ebbero iniziative simili anche a Milano, con la costruzione del calcolatore CRCI02A, e a Roma, con il FINAC.²⁶² Erano del resto gli anni di maggior sviluppo del programma cibernetico italiano (che trovava terreno fertile nel contemporaneo fiorire dell'informatica), che già a metà degli anni Sessanta del secolo scorso

²⁵⁸ Somenzi 1993: 178.

²⁵⁹ Corbellini, Cordeschi 2011: 33.

²⁶⁰ Salvia 2009: 506.

²⁶¹ Pogliano 2007: 93-94.

²⁶² Cuzzer 1993: 16-25.

poteva vantare diversi centri attivi, la cui specializzazione variava in base agli interessi di ricerca dei singoli promotori:

I centri di cibernetica sorti a Napoli con Caianiello e Braitenberg, a Genova con Borsellino e Gamba, a Milano con Ceccato e Maretti, si differenziavano tra loro nettamente, con prevalenza nel primo della neurofisiologia, nel secondo della biofisica e nel terzo della linguistica. Ancora più specializzati erano i centri dedicati al calcolo automatico, ai servomeccanismi e all'automazione industriale.²⁶³

Pur trattandosi di un tema sicuramente ostico per i filosofi italiani, il congresso ebbe una grande risonanza, con oltre quattrocento iscritti, due relazioni introduttive, più di settanta comunicazioni presentate all'apertura e decine d'interventi e repliche, andando ben oltre la cerchia degli addetti ai lavori e richiamando l'attenzione dei maggiori quotidiani, oltre che dell'allora unico canale esistente della RAI.

Nella sua relazione introduttiva, Somenzi giudicò anzitutto necessario ridefinire il concetto di "macchina" alla luce degli sviluppi recenti, in modo da evitare che "la discussione sui rapporti tra le macchine e i rispettivi inventori o utenti si estenda disordinatamente fino a comprendere ogni genere di strumento impiegato dall'uomo".²⁶⁴ Questa premessa aveva probabilmente lo scopo di confutare implicitamente la tesi promulgata nell'altra relazione introduttiva ai lavori al congresso, in cui Guzzo, a quel

²⁶³ Somenzi 1993: 166.

²⁶⁴ Somenzi 1967: 51.

tempo presidente della Società filosofica italiana, provava a recuperare il significato originale di alcune parole di uso comune. Inteso come *techne* il processo espressivo dell'intuizione, Guzzo definiva "tecnica" l'insieme delle norme applicate utilizzate durante il processo di comunicazione. Due aspetti solo apparentemente distinti che confluivano nella creazione artistica:

L'artista dice a tutti quel che dice a sé stesso, né può esprimersi senza comunicare a chi l'ascolta l'espressione del suo "mondo interiore". L'espressione si fa comunicazione: questa si rivela all'apparire della stessa espressione: *techne* e "tecnica" si congiungono.²⁶⁵

Essendo la tecnica (e di conseguenza la tecnologia) un aspetto della creatività umana ("non ci sono tecniche senza *techne*"²⁶⁶), i calcolatori non differivano molto, in linea teorica, da ciò che altri campi in apparenza diametralmente opposti, come l'arte e la religione, avevano prodotto. Così facendo, Guzzo negava alle macchine automatiche quell'estrema originalità che invece era intenzione di Somenzi sottolineare.²⁶⁷

Prima di tutto, egli riesaminava brevemente i mutamenti del concetto stesso di macchina dall'epoca della prima rivoluzione industriale in poi. Sarebbe prevalsa, per tutto l'Ottocento, una concezione della "macchina" che, da una parte, relegava al ruolo di semplice strumento tutti quegli oggetti che si

²⁶⁵ Guzzo 1967: 11.

²⁶⁶ Ivi.: 6.

²⁶⁷ Pogliano 2007: 108.

limitavano a rendere meglio utilizzabile l'energia muscolare dell'uomo ma che, dall'altra parte, esaltava le macchine "autonome" che necessitavano di energia, esattamente come gli uomini e gli animali. Questo parallelismo tra fisiologia meccanica e fisiologia umana portò alla distinzione concettuale tra "organo" e "funzione", permettendo ai materialisti una completa identificazione tra sistema nervoso e attività mentali, e ai dualisti la possibilità di riservare allo spirito umano "la fissazione e il riconoscimento delle funzioni che tali organi assolvono".²⁶⁸

Ma non sfuggì ad alcuni filosofi presenti al congresso l'esistenza di macchine che risparmiavano all'uomo un lavoro di tipo mentale, come l'orologio per la misurazione del tempo, o le calcolatrici meccaniche per quanto riguardava la computazione. La differenza sostanziale tra quest'ultime stava nel ruolo attivo svolto dall'uomo per il loro corretto funzionamento: fornitore di energia nel primo caso, d'informazioni nel secondo. La netta separazione tra questi due concetti sarebbe avvenuta solo a partire da metà Novecento, ma già Babbage aveva perseguito il piano di una calcolatrice universale, generalizzato poi "nella forma più astratta possibile" da Turing:

²⁶⁸ Somenzi 1967: 52.

Le macchine di Turing si prestano all'esecuzione non solo del lavoro deduttivo dell'uomo, ma anche di tutte le operazioni che caratterizzano i processi induttivi ed ogni altro processo mentale.²⁶⁹

Somenzi delineava quindi una prima definizione di macchina, comprendente tutte le macchine realizzate dall'uomo ma che, allo stesso tempo, escludeva tutte quelle tecniche e strumenti per cui “questa definizione suonerebbe impropria”:

La definizione pone l'uomo di fronte a una “macchina” ogni qual volta egli abbia a sua disposizione un “sistema suscettibile di assumere uno tra più stati diversi, una lista che associa ad ognuno di questi stati una particolare azione, e le regole per passare da uno stato ad un altro”.²⁷⁰

Per Somenzi, il filosofo non poteva più ignorare come il termine “macchina” comprendesse anche gli automi ideali, quelle macchine di Turing la cui realizzazione pratica non incontrava nessuna difficoltà di principio, utilizzate anche come modello matematico per spiegare l'attività mentale dell'uomo o per progettare sostituti artificiali del suo cervello sempre più efficienti. A riprova dell'importanza dell'argomento e della necessità di inquadrarlo sotto una luce nuova, vi erano migliaia di scritti pubblicati tra il 1950 e il 1965 che affrontavano il problema della possibilità che “una macchina pensi e che l'uomo sia una macchina di questo genere”. Non si trattava di date casuali: era del 1950 l'articolo di Turing *Computing*

²⁶⁹ Ivi.: 55.

²⁷⁰ Ibid.

Machinery and Intelligence, mentre risaliva al 1965 l'antologia di Somenzi che aveva presentato al pubblico italiano i principali articoli dedicati alla cibernetica.

Somenzi elencava poi rapidamente successi e insuccessi della cibernetica, ed è qui che si trova uno dei (pochissimi) riferimenti diretti all'AI fatti durante il congresso. Nell'intento di costruire macchine davvero intelligenti, l'uomo era costretto prima di tutto ad analizzare se stesso:

E' forse questo il motivo per cui le attività mentali dell'uomo [...] vengono ora analizzate con criteri diversi da specialisti di psicologia, pedagogia, neurofisiologia, biofisica, linguistica e logica matematica disposti a far convergere le proprie analisi verso conclusioni ben definite, e perciò utilizzabili dai costruttori di "intelligenze artificiali".²⁷¹

Oltre che ricondurre l'AI al più generale piano di ricerca interdisciplinare della cibernetica (negandole quindi lo statuto di scienza autonoma), in occasione del congresso Somenzi si dimostrò estremamente critico nei suoi riguardi. Tra gli insuccessi della cibernetica venivano citate le previsioni formulate da Newell e Simon nell'articolo del 1958 *Heuristic Problem Solving: The Next Advance in Operations Research*, giudicate "eccessivamente ottimistiche" considerando i risultati assai modesti²⁷²:

²⁷¹ Ivi.: 59.

²⁷² Per quanto il riferimento sia così chiaro da non lasciar dubbi, i due scienziati non venivano mai citati esplicitamente da Somenzi.

In pratica abbiamo assistito, fino ad oggi, solo alla dimostrazione di teoremi molto elementari con piccole varianti rispetto alle dimostrazioni note e [...] allo sviluppo di buone capacità nel campo dei giochi con regole molto ben definite e ad alcuni embrioni di conversazione tra macchina e uomo, nei quali l'uso di frasi fatte, compiuto per guadagnare nel "giuoco dell'imitazione", non può evitare che l'interlocutore si accorga presto di aver a che fare con una macchina e non con un minorato mentale.²⁷³

Dopo aver criticato (anche qui senza nessun riferimento esplicito) le prestazioni del *Logic Theorist* e le ricerche condotte da Samuel, Somenzi accusava gli studiosi di AI di rinunciare a ogni tentativo di definire il pensiero e il linguaggio, mirando alla costruzione diretta di macchine pensanti e parlanti. Cosa che comportava non pochi problemi, dato che una macchina costruita a partire da queste premesse non avrebbe mai potuto rinunciare all'apporto dell'uomo, e quindi non sarebbe mai stata "pensante" e "autonoma" nel vero senso della parola. Piuttosto che aggirare l'ostacolo principale (cioè una conoscenza lacunosa del sistema nervoso) ricorrendo a procedure che rispecchiavano solo superficialmente l'operare umano, sarebbe stato meglio far imitare alle macchine "l'unico modo di pensare e parlare in cui siamo esperti", cioè quello umano.

Ciò che emergeva dalla relazione di Somenzi era la sua totale aderenza alla teoria cibernetica, per la quale organismi e macchine potevano essere studiati sotto il profilo di una comune organizzazione funzionale. Di contro,

²⁷³ Somenzi 1967: 60.

egli era scettico su alcune posizioni del funzionalismo tipiche dell'AI classica. Così, programmi per calcolatore in cui la simulazione rinunciava del tutto alla riproduzione delle procedure effettivamente utilizzate da un organismo o dall'uomo nell'esecuzione di un compito, non erano da lui ritenuti interessanti per fini esplicativi, né particolarmente promettenti dal punto di vista produttivo. Come specificato nell'introduzione a *La filosofia degli automi*, l'indirizzo della cibernetica per lui più interessante, per quanto secondario nelle ricerche di Wiener, era (e sarebbe rimasto) quello della bionica, che egli concepiva come l'impresa in cui neurologi e biochimici, unitamente a informatici e ingegneri elettronici, cominciavano a sperimentare la riproduzione materiale artificiale delle strutture organiche e dei processi, individuali ed evolutivi.²⁷⁴ Era lo stesso Somenzi a confermare questa posizione in un'intervista condotta da Riccardo Urbani 30 anni dopo il congresso di Pisa:

Personalmente, io sono rimasto ancorato all'ideale di una imitazione dei cervelli naturali meno complessi dedicata non tanto alle loro prestazioni, quanto alla loro struttura, alle cui caratteristiche potrebbe avvicinarsi meglio un calcolatore chimico, sia pure lento e impreciso, che non un calcolatore elettronico.²⁷⁵

Per Somenzi, sarebbe stato meglio imitare i dispositivi "intelligenti" di esseri semplici, come le piante carnivore dotate di poche migliaia di neuroni, piuttosto che dedicarsi alla costruzione di "programmi capaci di simulare, e

²⁷⁴ Somenzi 1965: 9.

²⁷⁵ Urbani 1997: 513.

neppure male, il dialogo tra uno psicanalista ed un paziente”.²⁷⁶

Nel corso degli anni, egli avrebbe sempre mantenuto una certa distanza nei confronti dell'AI, pur cambiando almeno parzialmente il giudizio negativo affiorato durante il congresso. La terza edizione de *La filosofia degli automi* (datata 1986), ad esempio, curata questa volta insieme a Roberto Cordeschi, risultava notevolmente ampliata rispetto alle precedenti, e oltre ai “classici” della cibernetica includeva anche i più importanti articoli scritti da Minsky, McCarthy e Simon negli anni Cinquanta-Sessanta. Nella premessa al volume, Somenzi affermava di vedere nell'AI il proseguimento di almeno una parte del programma cibernetico originario, apprezzandone l'assunzione fisicalista (la mente come dispositivo fisico di elaborazione dell'informazione), pur giudicando la simulazione di funzioni umane su calcolatore non coerente con l'approccio bionico che gli stava più a cuore.²⁷⁷

L'unica altra relazione del congresso a trattare esplicitamente di AI fu quella di Evandro Agazzi dal titolo *Simulazione del pensiero e intelligenza artificiale*. Le idee qui esposte erano già presenti in un articolo (scritto in inglese) presentato durante il *Wiener Memorial Meeting on the Idea of Control* tenuto a Genova nel 1965, ma pubblicate (in italiano) solo nel 1967 sulla *Rivista di Filosofia Neoscolastica* con il titolo di *Alcune osservazioni sul problema dell'intelligenza artificiale*, di cui la relazione presentata al

²⁷⁶ Ibid. Somenzi si riferiva probabilmente al chatterbot *ELIZA*, sviluppato da Joseph Weizenbaum nel 1966.

²⁷⁷ Somenzi 1986: 7-8.

congresso era un riassunto.²⁷⁸

L'articolo di Agazzi conteneva una critica estensiva al problema dell'AI, insistendo particolarmente sull'intenzionalità quale discrimine decisivo tra intelligenza umana e artificiale, anticipando di circa 15 anni le idee di John Searle, ideatore del celebre esperimento delle stanze cinesi.²⁷⁹

Per Agazzi, l'obiettivo di costruire un artefatto in grado di pensare come gli esseri umani non era, in linea di principio, sbagliato, se l'assunto veniva considerato come una semplice idea-guida da cui partire nel tentativo di realizzare simulazioni ed emulazioni del pensiero umano. Questa ipotesi, tuttavia, non avrebbe potuto considerarsi scientificamente valida fino a quando non fosse stata sorretta da prove adeguate, non solo in termini di risultati, ma anche nelle procedure richieste per raggiungerli. Ma fino ad allora, sosteneva Agazzi, non era stata formulata nessuna prova conclusiva e rigorosa a riguardo: anche se certe macchine potevano ottenere risultati simili a quelli degli esseri umani, questo non comportava che il pensiero e la macchina fossero riconducibili alla stessa natura.

Agazzi sottolineava come all'origine di questi fraintendimenti vi fosse l'antropomorfizzazione linguistica delle operazioni eseguite dalle macchine,

²⁷⁸ Bianca 2015: 92.

²⁷⁹ Searle immaginava una situazione in cui una persona, chiusa in una stanza, riceveva domande scritte in cinese (cioè una lingua che non conosceva), ma con la possibilità di farvi corrispondere risposte corrette (anch'esse in cinese) usando un manuale scritto in una lingua conosciuta. Ma in questo caso una risposta esatta alla domanda posta non comportava una comprensione autentica del loro contenuto. Questo, osservava Searle, accadeva anche con i computer e con tutti i sistemi computazionali, a riprova del fatto che la semplice manipolazione di simboli non avrebbe mai permesso la completa comprensione dei significati. Lo scopo di Searle era quello di criticare i presupposti della cosiddetta AI "forte" (termine da lui stesso coniato), secondo cui un computer opportunamente programmato corrisponderebbe in tutto e per tutta a una "mente". Cfr. Searle 1980.

come quando si affermava che un calcolatore elettronico potesse “imparare”, “ricordare” o “fare delle scelte”. L'uso di questi termini, appartenenti al linguaggio comune, portava a credere che la macchina eseguisse le stesse operazioni svolte dalla mente umana. Secondo Agazzi, questa pratica linguistica era giustificata solo quando riguardava le operazioni transitive, cioè quel tipo di operazioni definite sulla base di un risultato che qualsiasi agente (sia esso un uomo o una macchina) poteva raggiungere partendo da uno stato iniziale: “cucire”, ad esempio, rientrava in quest’ambito. Nel caso di espressioni come “rispondere”, “apprendere”, “ricordare” ecc., esse non erano operazioni transitive, ma immanenti, cioè non definibili da risultati “esterni”, ma dalle modifiche indotte “all’interno” dell’agente. Lo stesso test di Turing veniva giudicato da Agazzi un criterio insufficiente per stabilire l’equivalenza semantica tra i predicati psicologici usati per l’uomo e per gli automi, in quanto criterio di natura “comportamentistica”.

Nei decenni che seguirono il congresso Agazzi si sarebbe dedicato spesso a questi problemi: in tal senso, *Alcune osservazioni sul problema dell'intelligenza artificiale e Operazionalità e intenzionalità: l'anello mancante dell'intelligenza artificiale* possono essere visti come il punto di partenza e di arrivo di un discorso iniziato nel 1967 e giunto alla sua naturale conclusione nel 1991. Se il primo articolo conteneva già i punti fondamentali della sua riflessione sull'AI, quello del 1991 sottolineava la stretta correlazione tra intenzionalità e operazionalità, presupposto su cui si

basava tutta la riflessione sviluppata da Agazzi sulla conoscenza umana: un collegamento la cui importanza era ben esemplificata nelle sue considerazioni sull'AI.²⁸⁰

Come sottolineato da Urbani, l'atteggiamento generale dei filosofi nostrani durante il congresso fu di sostanziale riluttanza a cogliere i nuovi contenuti dell'analogia uomo-macchina. Ciò che mancava era una critica strutturale e specifica ai temi esposti nella relazione di Somenzi: non fu ignorata solo l'AI, ma anche la cibernetica nel suo complesso, in favore di un'analisi del rapporto uomo-macchina il più generale possibile che preferiva citare, ad esempio, Platone piuttosto che Shannon. Una delle poche eccezioni fu il test di Turing, ripreso più volte nel corso degli interventi, pur con il solo scopo di confutarlo. Mentre i partecipanti erano generalmente propensi a parlare dell'essere umano in termini meccanicistici per quanto riguardava le sue facoltà biologiche inferiori, molti non erano disposti a fare altrettanto per quelle superiori, come il linguaggio e l'autocoscienza, e ciò si traduceva nel rifiuto di considerare il gioco dell'imitazione (in cui quelle questioni perdevano di significato) un criterio valido. Molti tendevano a fare propria la posizione espressa da Guido Calogero in un saggio del 1965 e ribadita da lui stesso durante il congresso²⁸¹:

²⁸⁰ Bianca 2015: 92.

²⁸¹ Calogero 1966: 1-11.

Una macchina non “dialoga” se unicamente risponde a domande: “dialoga” soltanto se ha interesse a che l’interrogante sia soddisfatto della sua risposta.²⁸²

Esattamente come l’orologio (esempio citato da Somenzi, che a sua volta era ripreso da La Mattrie), la macchina si limitava a produrre certi fenomeni che potevano essere interpretati in un certo modo, ma che non presupponevano nessuna forma di coscienza. Un automa capace di risolvere, ad esempio, delle equazioni algebriche, ma senza conoscerne il risultato e gli effetti, veniva paragonato a un essere umano che mentre leggeva una pagina ad alta voce pensava ad altro, non prestando attenzione al suo significato. Per Calogero, l’uomo possedeva un lato “automatesco” che lo rendeva simile alla macchina, grazie a cui era capace di svolgere certe funzioni senza pensare; di contro, egli possedeva anche una coscienza, la quale implicava “sempre piacere o dolore, cose che stanno al di là del regno della matematica e dei sillogismi”.²⁸³ Era solo l’incontro tra due individui dotati di coscienza a rendere possibile il riconoscimento di un altro essere: L’“altro” è colui che può soffrire, non colui che ragiona. [...] Un automa può risolvere problemi di calcolo infinitesimale, e non essere un “altro”, se il risolvere quei problemi non gli dà né gioia né noia.²⁸⁴

Per Calogero, la sola indagine significativa da condurre sugli automi avrebbe dovuto essere di tipo morale, ponendo l’accento sui loro

²⁸² Calogero 1967: 95.

²⁸³ Ivi.: 101.

²⁸⁴ Ibid.

“sentimenti” piuttosto che sulla loro intelligenza.

Alla luce di tutto ciò, non sorprendono le parole di Somenzi, che a fine lavori sostenne che forse sarebbe stato meglio organizzare un vero e proprio congresso di cibernetica seguito da una discussione filosofica specifica, che trattasse solo l’aspetto cibernetico dei rapporti tra uomo e macchina.²⁸⁵

Prima, durante e dopo l’iniziativa, lo stesso Somenzi raccolse i vari contributi al congresso, pubblicati in appendice al terzo volume degli Atti.

Un po’ come nel caso di *Adamo II*, anche qui le reazioni oscillavano dalla cieca fiducia nei confronti del progresso scientifico alla paura di

un’ipotetica distopia meccanica: “Scienza e filosofia verso un’intesa”²⁸⁶,

“La macchina farà più libero l’uomo”²⁸⁷, “La Cibernetica non spiega

tutto”²⁸⁸, “Il progresso tecnologico è beneficio solo se è controllato da un vigile senso morale”²⁸⁹.

Del congresso pisano diede una valutazione molto critica *Civiltà delle macchine*, sostenendo che da potenziale incontro costruttivo tra opinioni contrastanti si era invece trasformato in un sovrapporsi di voci inascoltate (e con poca voglia di ascoltare). Un po’ come Somenzi (ma con un tono molto più duro), *Civiltà delle macchine* rimproverava agli organizzatori e ai partecipanti il non aver stabilito fin dal principio un tema più specifico, una

²⁸⁵ Somenzi 1967: 287.

²⁸⁶ *Il Telegrafo*, 22 aprile 1967.

²⁸⁷ *Avanti!*, 4 maggio 1967.

²⁸⁸ *La Fiera letteraria*, 11 maggio 1967.

²⁸⁹ *Il Popolo*, 6 maggio 1967.

piattaforma tematica e linguistica comune, in modo da permettere l'incontro tra i due diversi schieramenti:

La conseguenza di queste babeliche contraddizioni è stata che un linguaggio ibrido, aporetico e non analizzato, ordito di anacoluti concettuali e di equivoca emotività, è stato concordemente impiegato come uno strumento linguistico interdisciplinare, di cui invece e a ragione Hume avrebbe detto: “Gettalo nel fuoco, perché non può contenere altro che dei sofismi e delle illusioni”.²⁹⁰

Meno severo fu il giudizio di altre riviste specialistiche, come *De Homine* e *Ricerche metodologiche*, che comunque chiesero una maggiore collaborazione tra scienza e filosofia, in modo da superare quegli “opposti estremismi” manifestati durante il congresso.²⁹¹

Nello stesso anno si tenne un altro evento, forse meno d'impatto sulla cultura italiana, ma comunque di un certo interesse per quanto riguarda lo sviluppo dell'AI nel nostro paese. Dal 16 al 19 ottobre 1967 trentaquattro relatori di svariata appartenenza disciplinare vennero infatti convocati a Roma dall'Accademia dei Lincei per esporre il loro punto di vista sulle implicazioni scientifiche, tecniche e sociali dell'automazione elettronica. L'allora presidente Beniamino Segre aprì le sedute osservando come fossero diversissimi i pareri sull'importanza, anche “spirituale”, delle calcolatrici elettroniche; una diversità che qualche mese prima si era vista negli animati dibattiti del XXI Congresso nazionale dei filosofi tra i

²⁹⁰ i.g. 1967: 2.

²⁹¹ Urbani 1997: 511.

sostenitori dell'“Era Cibernetica” e chi profetizzava il predominio della macchina sull'uomo. A fine discorso, Segre augurava che l'incontro tra i partecipanti potesse promuovere scambi d'idee e confronti proficui, il tutto a beneficio di quella nuova disciplina (l'informatica) che attestava “le sbalorditive possibilità della tecnica e dell'ingegno umano”.²⁹²

A differenza di quanto avvenuto a Pisa, il congresso su *L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali* fu caratterizzato da un maggior “tecnicismo” nelle relazioni, oltre che da una maggiore consapevolezza degli argomenti trattati (e non poteva essere altrimenti, visti i nomi dei partecipanti).

Il matematico Alessandro Faedo (rettore a Pisa), ad esempio, illustrò *l'Origine ed evoluzione dei calcolatori automatici* citando, tra le altre cose, la tecnica del *time-sharing* inventata quasi per caso da McCarthy nel 1958.²⁹³ Gianfranco Capriz, in quegli anni direttore del CSCE (Centro Studi Calcolatrici Elettroniche) e professore di meccanica razionale a Pisa²⁹⁴, parlò della necessità di introdurre “procedure euristiche” nella ricerca sistematica, citando i risultati “impressionanti” del *Logic Theorist* e della macchina prova-teoria di David Gelertner:

²⁹² Segre 1968: 12.

²⁹³ L'espressione *time-sharing* (spesso resa in italiano con “divisione di tempo” o, talvolta, con “crono ripartizione”) indica il funzionamento “a multiprogrammazione” che consente a più utenti di usufruire contemporaneamente un calcolatore, pur utilizzando programmi diversi. Questa tecnica ha il vantaggio di sfruttare tutta la potenza del processore centrale e sopperire così alla lentezza dei dispositivi di elaborazione degli input e degli output. Per quanto non direttamente collegato con l'AI, il *time-sharing* può essere visto come una sua diretta conseguenza. L'intento di McCarthy era quello di consentire ai ricercatori l'utilizzo del calcolatore ogni qual volta fosse necessario, dal momento che il modo migliore per testare una teoria era “dare al calcolatore un comando e vedere cosa succedeva”. Cfr. McCorduck 1987: 217.

²⁹⁴ Oggi è professore emerito.

Il procedimento usato in questo ultimo caso è interessante: assegnata una proposizione della quale si vuole stabilire la verità, si procede a tracciare un diagramma rappresentativo ma con dati inessenziali scelti a caso. Il programma esplora il diagramma alla ricerca di segmenti paralleli, o di eguale lunghezza, ecc. e controlla poi se le relazioni trovate siano puramente casuali o non possano invece servire per la dimostrazione.²⁹⁵

Durante il secondo giorno del congresso Giuseppe Montalenti aprì una seduta sulle applicazioni dei calcolatori elettronici nell'ambito della biologia, le quali comprendevano anche lo studio diretto dei fenomeni elementari del sistema nervoso per mezzo di modelli computerizzati:

Le funzioni del sistema nervoso, le funzioni superiori del pensiero, dell'intelligenza, oggi sono accessibili ad una indagine meccanicistica, appunto grazie all'uso dei calcolatori elettronici.²⁹⁶

La questione venne approfondita da Valentino Braitenberg, che parlò del rinnovato slancio della “scienza dei cervelli” avvenuto grazie all'avvento della cibernetica (“siamo liberi di considerare il cervello alla stregua di una macchina essenzialmente analogica”²⁹⁷), e da Antonio Borsellino, che preferì soffermarsi su alcune delle implicazioni filosofiche poste dalla costruzione di macchina “intelligenti”. Borsellino richiamava la definizione di intelligenza concettuale così com'era stata definita da Piaget, cioè di

²⁹⁵ Capriz 1968: 57.

²⁹⁶ Montalenti 1968:163.

²⁹⁷ Borsellino 1968: 215.

un'attività reversibile, operante su o con un linguaggio, inteso come “comportamento potenziale”, simbolico, svincolato dal presente percettivo:

Nulla sappiamo ancora oggi del modo come il patrimonio linguistico, nelle sue varie modalità, sia acquisito, codificato, conservato, richiamato e, tradotto in attività di neuroni, messo in circolazione nei vari canali ed utilizzato nell'attività “simbolica”, costituente l'intelligenza concettuale.²⁹⁸

Tutti problemi che, secondo Borsellino, avevano comportato il sostanziale fallimento tecnico delle “macchine intelligenti”, dovuto probabilmente all'ambizione di voler fare troppo con troppo poco; per quanto estremamente interessanti, i calcolatori fino a quel momento costruiti disponevano di troppo pochi sensori d'ingresso e di memorie troppo limitate per poter competere con l'enorme ricchezza degli input sensoriali e con la complessità dei livelli elaborativi delle intelligenze naturali.

A differenza di quanto avvenuto a Pisa, Somenzi decise di dedicare il proprio intervento esclusivamente agli *Aspetti filosofici del problema della “intelligenza artificiale”*. Ricorrendo alla definizione di AI formulata da Feigenbaum e Feldman nel 1963 (“la costruzione di programmi per calcolatori, i quali esibiscano il comportamento che noi chiamiamo comportamento intelligente quando lo osserviamo negli esseri umani”²⁹⁹),

Somenzi scartava a priori, “se non altro per ragioni di prudenza”, che il

²⁹⁸ Ibid.: 225.

²⁹⁹ Feigenbaum, Feldman 1963: 3. Si noti come la definizione fosse molto simile a quella formulata da McCarthy nella Proposta di Dartmouth del 1955. Tuttavia, trattandosi di un documento a quel tempo riservato, è improbabile che Somenzi lo conoscesse: da qui la citazione “alternativa” piuttosto che quella originale.

pensiero umano fosse qualcosa di non analizzabile esaurientemente in operazioni ripetibili da macchine. Nonostante molti scienziati si fossero occupati della questione, restava ancora diffuso il timore che da ulteriori successi dei calcolatori nella simulazione delle attività mentali venisse tratta una serie di conseguenze “materialistiche”. Un primo atteggiamento, sosteneva Somenzi, consisteva nell’isolare il campo dell’intelligenza artificiale (intesa non come disciplina, ma come comportamento esibito dalla macchina) da quello dell’intelligenza naturale, considerandoli incomparabili; egli confutava questa posizione illustrando il metodo tipico dei tecnici dell’AI, i quali miravano a imitare i soli risultati dell’attività mentale mediante materiali che non avevano nulla a che fare con quelli costituenti il cervello. Ma nulla escludeva che dai progressi nel settore dell’AI potessero derivarne altri nella comprensione del modo in cui il cervello umano “pensa”:

Per esempio, dalla costruzione di complicati programmi per la risoluzione automatica di problemi da parte dei calcolatori, Newell, Simon e collaboratori hanno ricavato nuovi efficaci strumenti per la costruzione di teorie sul pensiero umano.³⁰⁰

Un secondo tipo di argomentazione consisteva nel sottolineare le limitazioni dei calcolatori rispetto all’intelligenza umana; intelligenza i cui confini venivano spostati, di volta in volta, sempre più in alto dai pessimisti,

³⁰⁰ Somenzi 1968: 229.

in modo che non vi rientrassero quei fenomeni di cui si era raggiunta una spiegazione sufficiente o la ripetizione da parte di meccanismi. Già Oettinger, McCulloch e von Neumann avevano confutato questa posizione tramite una specie di “teoria di esistenza”: esisteva almeno una soluzione per il problema di ottenere ogni prestazione intelligente mediante un apparato materiale limitato nel peso, nel volume e nel consumo di energia. Questo, ovviamente, comportava la “tanto contestata” identificazione tra pensiero e attività della materia cerebrale: ma lo studio dei rapporti tra attività cerebrale e mentale non richiedeva che il calcolatore stesso fosse costruito secondo tale teoria. Era esclusivamente per questo motivo che i metodi di realizzazione dell’AI e gli studi tradizionali sull’intelligenza naturale differivano così tanto: la contestazione, insomma, poggiava su premesse vuote.

Riguardo all’uso della terminologia antropomorfa nella descrizione dei calcolatori, Somenzi riprendeva e faceva sua la tesi sostenuta da Agazzi durante il congresso di Pisa, ribaltandone però le conclusioni pessimistiche circa il futuro del pensiero artificiale:

Un cervello artificiale davvero fedele al suo modello umano potrebbe domani svolgerne tutte le attività, ivi compresa quella consistente nell’invenzione tipicamente antropocentrica di un “io” distinto dal cervello stesso, il quale ritenga

di stare adoperando un organo come una specie di protesi naturale concessagli per pensare, ma insufficiente da sola a svolgere la benché minima attività mentale.³⁰¹

Poche settimane dopo il convegno linceo, Italo Calvino iniziò un giro di conferenze scegliendo di parlare di *Cibernetica e fantasmi*. Secondo Calvino, la cultura moderna stava iniziando a percepire il mondo sempre più come un tutto discreto (ovvero composto da parti separate) piuttosto che continuo. Ma non era sempre stato così:

Il pensiero, che fino a ieri ci appariva come qualcosa di fluido, evocava in noi immagini lineari come un fiume che scorre o un filo che si dipana, oppure immagini gassose, come una specie di nuvola, tant'è vero che veniva spesso chiamato "lo spirito", – oggi tendiamo a vederlo come una serie di stati discontinui, di combinazioni di impulsi su un numero finito (un numero enorme ma finito) di organi sensori e di controllo.³⁰²

Dalla consapevolezza dell'esistenza di queste combinazioni nasceva anche una seconda riflessione, quella legata alla tematica dei cervelli elettronici, il cui limite era sicuramente quello di non poter riprodurre tutte le funzioni del cervello umano, ma che potevano rivelarsi utili come modello teorico convincente per i processi più complessi della memoria umana, delle associazioni mentali, dell'immaginazione e della coscienza.

Tra le personalità che avevano determinato un radicale cambiamento nei

³⁰¹ Ivi.: 232.

³⁰² Calvino 1967: 209.

processi mentali umani, Calvino individuava Wiener, Shannon e von Neumann:

Al posto di quella nuvola cangiante che portavamo nella testa fino a ieri e del cui addensarsi o disperdersi cercavamo di renderci conto descrivendo impalpabili stati psicologici, umbratili paesaggi dell'anima, – al posto di questi oggi sentiamo il velocissimo passaggio di segnali sugli intricati circuiti che collegano i relè, i diodi, i transistor, di cui la nostra calotta cranica è stipata. Sappiamo che, come nessun giocatore di scacchi potrà vivere abbastanza a lungo per esaurire le combinazioni delle possibili mosse dei trentadue pezzi sulla scacchiera, così dato che la nostra mente è una scacchiera in cui sono messi in gioco centinaia di migliaia di pezzi – neppure in una vita che durasse quanto l'universo si arriverebbe a giocare tutte le partite possibili.³⁰³

L'influsso di Von Neumann e degli altri aveva, dunque, fornito nuove consapevolezze: l'arte combinatoria nasceva e si esprimeva come processo rivolto a una rivincita della discontinuità, divisibilità, combinatorietà su tutto ciò che costituisce corso continuo.³⁰⁴

³⁰³ Calvino 1980: 167-168.

³⁰⁴ Buffo 2011.

5. AI@50

Nel 2006, in occasione del cinquantesimo anniversario del *Dartmouth Summer Research Project on Artificial Intelligence*, più di 100 studenti e ricercatori si riunirono a Dartmouth per la *Dartmouth Artificial Intelligence Conference: The Next Fifty Years* (o *AI@50*). Erano tre gli scopi di questa conferenza: celebrare quella del 1956; discutere sui progressi dell'AI; e stabilire quali sarebbero stati o avrebbero dovuto esserne gli sviluppi futuri. James Moor, professore di filosofia al Dartmouth College e organizzatore della conferenza, affermò che lo scopo dei ricercatori riunitisi 50 anni prima ad Hannover era stato quello di creare macchine più “consapevoli”, e sulla base di quest’obiettivo tracciare un quadro di riferimento su cosa fosse effettivamente l’intelligenza umana. Riguardo la conferenza del 2006, egli sottolineava:

We expect to undertake a full exploration into the many emerging directions for future AI research, just as the College took the first steps to establish AI as a research discipline 50 years ago.³⁰⁵

Secondo Carol Folt, preside della Faculty of Arts and Sciences e professoressa di scienze biologiche a Dartmouth, era più che comprensibile che l’AI, come campo di ricerca, attirasse le menti più brillanti e fantasiose, coloro che sono soliti lavorare senza tener conto dei confini disciplinari, dal momento che l’innovazione e l’interdisciplinarietà erano i due principali

³⁰⁵ Knapp 2006c.

tratti distintivi della conferenza originale.³⁰⁶

Barry Scherr, a quel tempo rettore, aggiunse che il successo del *workshop* del 1956 risiedeva nello spirito che aveva generato e che, al di là della semplice commemorazione, era proprio questo aspetto che gli organizzatori avevano intenzione di replicare:

The continuing accomplishments in the years since have proven that the field of AI remains vital and filled with promise. I hope that the *AI@50* participants enjoyed recalling the early years of AI at the same time that they were helping to develop a road map for future avenues of study.³⁰⁷

Sempre secondo Scherr, proprio perché l'AI era andata, nel corso degli anni, sempre più legandosi alla società, dalla vita di tutti i giorni fino agli aspetti più globali, era necessario interrogarsi circa la direzione attuale della ricerca, sul dove si sarebbe arrivati, sul come sarebbero stati raggiunti questi risultati e quali implicazioni etiche avrebbe avuto tutto il processo.

Tenutasi dal 13 al 15 luglio 2006, durante la conferenza furono presenti cinque dei partecipanti del congresso originale: Marvin Minsky, Ray Solomonoff, Oliver Selfridge, Trenchard More e John McCarthy. Rodney Brooks, direttore del laboratorio d'informatica e intelligenza artificiale del Computer Science and Artificial Intelligence Lab (CSAIL), sottolineava il carattere quasi “sovversivo” di questi primi ricercatori, specie considerando le implicazioni filosofiche insite nel concetto di AI:

³⁰⁶ Knapp 2006a.

³⁰⁷ Ibid.

At the 1956 Dartmouth Artificial Intelligence Conference an audacious, outrageous even, intellectual Zeitgeist emerged: that the core of humanity, our ability to think and reason, was subject to our own technological understanding, a recursive formulation of our very nature. And the participants were right.³⁰⁸

La conferenza fu finanziata dall'Office of the Dean of the Faculty e dall'Office of the Provost di Dartmouth, oltre che dalla Frederick Whittemore Foundation, dalla General Electric Foundation e in particolare dalla DARPA (Defense Advanced Research Projects Agency), da cui ricevette 200,000 dollari.

Intervistato a riguardo, il commento di Moor fu che il contributo della DARPA avrebbe permesso di guardare al futuro dell'AI, con particolare attenzione verso quei modelli e quelle metodologie che avrebbero potuto portare dei progressi (e dei successi) in aree di ricerca specifiche quali l'apprendimento, la vista, il ragionamento, la ricerca, il linguaggio naturale, e la cognizione.³⁰⁹ Secondo Moor, nell'odierno mondo tecnologizzato, un progresso così inteso aveva portato a risultati quali intelligenze artificiali in grado di giocare (e vincere) a scacchi, programmi in grado di tradurre la lingua araba e di raccogliere informazioni vitali per la sicurezza nazionale. Oltre all'aver fornito questo fondamentale supporto per l'AI@50, Moor sostenne che il finanziamento della DARPA avrebbe permesso a 25 laureandi e post laureandi di unirsi ai 31 scienziati di fama internazionale

³⁰⁸ Knapp 2006c.

³⁰⁹ Knapp 2006b.

nel campo dell'AI invitati per la conferenza. L'idea di far partecipare questi giovani ricercatori venne da George Cybenko, professore d'ingegneria a Dartmouth oltre che principale responsabile per i finanziamenti della DARPA: questa inclusione gli avrebbe permesso non solo di avviare la loro carriera lavorativa, ma ciò rientrava, secondo Cybenko, anche nello spirito della conferenza originale, quando “those first-ever discussions were driven by equally young researchers who had an enormous vested interest in the future of AI”.³¹⁰

Per l'AI@50, la DARPA richiese una relazione riassuntiva tratta dagli atti del convegno, la quale avrebbe dovuto: analizzare i progressi della “sfida originale” dell'AI durante questi primi cinquant'anni, stabilendo se essa si fosse rivelata più facile o più difficile del previsto, motivando la risposta; documentare quali, secondo i partecipanti della conferenza, sarebbero state le principali sfide nel campo della metodologia e dello sviluppo, e quali conquiste sarebbero state necessarie per il loro superamento; infine, confrontare questi obiettivi e risultati con gli sviluppi e le tendenze delle altre aree di ricerca, come la teoria del controllo, quella dell'elaborazione dei dati, la statistica e la teoria dell'informazione.

La DARPA chiese anche ai partecipanti di concentrarsi sulle esigenze della difesa e della sicurezza nazionale degli Stati Uniti. Durante la conferenza, Charles Holland, direttore del DARPA Information Processing Technology Office, condusse una tavola rotonda sul ruolo dell'AI nello sviluppo della

³¹⁰ Ibid.

futura tecnologia intelligente.

Per evitare malintesi di natura politica, nel proprio sito web il Dartmouth College precisa che “DARPA's support for AI@50 does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred”.³¹¹

3.1 McCarthy e Minsky sullo stato dell'AI

Il titolo del primo intervento della conferenza fu *Tonypandy and the Origin of a Science*, a opera di Carey Heckman, Senior Lecturer dell'Università di Dartmouth.

Heckman si focalizzò sulla storia della conferenza del 1956. *Tonypandy* significa “saggezza acquisita”, storie che si presuppone siano vere, quasi leggende: il termine migliore per descrivere la riunione originale. Non si trattò infatti di una conferenza vera e propria, ma di un insieme di “glorified poster sessions” durate 8 settimane.³¹² Heckman sottolineava anche come il termine “artificial intelligence” venisse già coniato nella Proposta di Dartmouth, non durante la conferenza, e come nel periodo di tempo che intercorse tra la proposta e il congresso vero e proprio esso fosse già diventato un termine abbastanza ricorrente.³¹³ Il più grande risultato del *Dartmouth Summer Research Project on Artificial Intelligence* restava,

³¹¹ Ibid.

³¹² Houston 2006d.

³¹³ In realtà, come si è visto, è probabile che il termine circolasse già da prima della presentazione della proposta.

comunque, quello di aver riunito alcuni dei migliori scienziati del tempo, ognuno impegnato in un diverso percorso di ricerca, all'interno di un progetto generale incentrato sulla simulazione del pensiero umano.

Curiosamente, il più significativo contributo ai posteri lasciato dal progetto non fu il resoconto della conferenza (che non venne mai redatto), ma la proposta stessa.

Nell'intervento immediatamente successivo, *What Was Expected, What We Did, and AI Today*, McCarthy esordì subito correggendo Heckman:

“Artificial intelligence is not, by definition, simulation of human intelligence”.³¹⁴

Il ruolo simbolico della Dartmouth Conference fu più importante dei suoi effetti immediati. L'idea originale era quella che i partecipanti avrebbero dovuto collaborare, ma ognuno era talmente concentrato sul proprio programma di ricerca da prestare ben poca attenzione al lavoro degli altri.

Ma, oltre ai problemi organizzativi e personali, a ridimensionare l'entusiasmo generale fu la consapevolezza, sopraggiunta dopo la conferenza, che occuparsi dell'AI si sarebbe rivelato molto più difficile del previsto.³¹⁵

Una delle domande del pubblico fu: “When will we have human-level AI?”.

Per McCarthy, una domanda del genere era priva di senso. Una domanda adeguatamente posta avrebbe dovuto concentrarsi sull'idea che un'AI di

³¹⁴ Houston 2006b.

³¹⁵ Ibid.

livello umano sarebbe stata possibile solo risolvendo alcuni dei suoi problemi di base. Erano tre i casi specifici menzionati da McCarthy. Il primo di essi era il cosiddetto *frame problem*, il quale insorge quando si prova a formalizzare un processo di *problem-solving* che coinvolge le interazioni con un mondo complesso. Più nello specifico, riguarda la difficoltà nel mantenere le tracce delle conseguenze di un'azione compiuta in una rappresentazione del mondo.³¹⁶ McCarthy, insieme a Patrick J. Hayes, si occupò di questo problema già nel 1969 nell'articolo *Some Philosophical Problems from the Standpoint of Artificial Intelligence*:

If we had a number of actions to be performed in sequence we would have quite a number of conditions to write down that certain actions do not change the values of certain fluents. In fact with n actions and m fluents we might have to write down mn such conditions. [...] A number of fluents are declared as attached to the frame and the effect of an action is described by telling which fluents are changed, all others being presumed unchanged.³¹⁷

Per risolvere il problema bisognava evitare di specificare cosa non avviene quando un'azione viene compiuta.

Il secondo problema era quello della “qualificazione” (*qualification problem*): esso riguarda l'impossibilità di assumere tutte le precondizioni necessarie a un'azione nel mondo reale per ottenere l'effetto voluto. Anche in questo caso McCarthy si occupò della questione già molti anni prima:

³¹⁶ Hayes 1981.

³¹⁷ Hayes, McCarthy 1969: 31.

For example, the successful use of a boat to cross a river requires, if the boat is a rowboat, that the oars and rowlocks be present and unbroken, and that they fit each other. Many other qualifications can be added, making the rules for using a rowboat almost impossible to apply, and yet anyone will still be able to think of additional requirements not yet stated.³¹⁸

Sempre secondo McCarthy, una possibile soluzione del problema consisterebbe nel soffermarsi non sulle precondizioni che rendono possibile un'azione, ma in quelle che ne ostacolerebbero il successo: "If no lack of oars or other circumstance preventing boat use is deducible, then the boat is concluded to be usable".³¹⁹

L'ultimo era il problema della ramificazione (*ramification problem*), ovvero come evitare di rappresentare tutte le conseguenze implicite di una determinata azione.

Durante l'*AI@50*, McCarthy affermò che "all three have been solved in important contexts and for important applications, but none have been solved at the human level of intelligence"³²⁰: per essere sviluppate al meglio, queste idee richiedevano un'estensione della logica, principalmente di quella non-monotona. Egli avanzava anche la possibilità che esistessero altri problemi oltre quelli citati, sebbene fino a quel momento nessuno li avesse scoperti.

Provando, in definitiva, a dare una risposta alla domanda iniziale, McCarthy

³¹⁸ McCarthy 1980.

³¹⁹ Ibid.

³²⁰ Houston 2006b.

afferitava che, per raggiungere un'intelligenza artificiale paragonabile a quella umana, sarebbe stata necessaria una sorta di autoconsapevolezza (*self-awareness*) artificiale.

Un'altra domanda del pubblico fu: "Where are we in simulating human-year reasoning? Like a 1-year-old? A 2-year-old?".³²¹ McCarthy rispose che per certi aspetti le macchine erano incredibilmente avanzate, mentre per altri non erano in grado di raggiungere la maturità di un bambino al primo anno di vita. Chiudendo con una nota d'amarezza:

One of my complaints is that people in AI, psychology, and philosophy are too apt to focus on appearance rather than the reality that appears behind the appearance. Just drawing patterns of appearance (of things) is not enough. We are middle-sized objects, and we have the ability to recognize other middle-sized objects that existed before us.³²²

All'intervento di McCarthy seguì quello di Minsky, *The Emotion Machine*, il cui titolo riprendeva quello del suo ultimo libro, presentato proprio durante la conferenza.

Minsky esordì con un aneddoto: stava guidando in giro per la città con a bordo la figlia di 18 mesi, la quale urlava periodicamente "Attenzione!" (*care*). Minsky e la moglie non riuscivano a capire a cosa si riferisse. Scoprirono poi che la figlia aveva visto uno spot pubblicitario dell'organizzazione CARE, il cui logo imitava la tecnica dello stencil.

³²¹ Ibid.

³²² Ibid.

Durante il giro, le era capitato di notare pali del telefono con numeri impressi, e aveva associato quel particolare tipo di incisione alla parola *care*.

La morale di fondo: “Logic and reasoning are complex!”.³²³

Molti dei primi ricercatori dell’AI, negli anni Sessanta e Settanta, risolvevano problemi di logica e matematica di livello liceale o universitario, perché era quello il loro campo di competenza iniziale. A distanza di decenni, le cose non sembravano cambiate molto: non si era ancora riusciti a costruire un programma visivo in grado di riconoscere gli oggetti di una normalissima stanza, o un programma che tramite il linguaggio naturale avrebbe potuto rispondere a domande su argomenti semplici, come quelle che fa un bambino quando ascolta una fiaba. Per Minsky molti dei moderni ricercatori dell’AI soffrivano di “invidia nei confronti della fisica” (*physics envy*): desidererebbero principi formali generali da usare come teoria del pensiero. A questo riguardo il commento di Minsky fu: “AI-ers use Occam's Razor too much!”.³²⁴ Il cervello è costituito da miriadi di architetture, ed è in grado di usare procedure e rappresentazioni molto diverse tra di loro: la biologia è disordinata, non un tutto omogeneo. Per capire davvero un fenomeno cognitivo, servivano almeno 3 diverse teorie a riguardo, perché è probabile che il cervello lo produca utilizzando contemporaneamente molti metodi diversi tra loro. Secondo Minsky, le teorie multiple erano sempre preferibili a quelle

³²³ Ibid.

³²⁴ Ibid.

complesse, eleganti e rigidamente unitarie. La mente veniva vista come un grande puzzle, in cui ogni pezzo poteva rappresentare una teoria matematica: ma questi pezzi andavano collegati tra di loro per formare il quadro d'insieme. Non esiste un solo modo per affrontare un problema, e sono processi come l'analogia, la progettazione, la semplificazione, la riformulazione e la contraddizione a rendere la specie umana intelligente. Bisognava, in definitiva, elaborare una macchina in grado di riflettere davvero: erano questi i principali temi esposti in *The Emotion Machine*.

3.2 Uno sguardo sul presente

Furono più di 30 le presentazioni e gli eventi dell'*AI@50* che ebbero come oggetto l'apprendimento, la ricerca, la robotica, il linguaggio e la cognizione: essi documentavano i significativi successi ottenuti dall'AI nel corso degli ultimi 50 anni. La robotica ne è un esempio perfetto.

Nel suo discorso *Making Bodies Smart*, all'interno della conferenza sulle *Future Interactions with Intelligent Machines*, Daniela Rus, direttrice del CSAIL (Computer Science and Artificial Intelligence Laboratory) al MIT, definiva la robotica come il sogno di costruire macchine a nostra immagine e somiglianza, al fine di studiare la vita e l'intelligenza, oltre che le connessione tra percezione e azione.

Dal suo punto di vista, la storia della robotica coincideva con la storia degli automi, dai Golem a Frankenstein fino ai robot attuali; questi ultimi tuttavia

erano molto diversi dagli automi fissi degli anni Cinquanta, in grado di compiere solo pochi compiti molto specifici. Nel 2006, i robot erano ormai ovunque e in grado di eseguire tantissimi compiti: pulivano case, esploravano la superficie di Marte e il fondo degli oceani, e vincevano la DARPA Grand Challenge in una corsa di 132 miglia nel deserto del Mojave. Rus ipotizzava che nel futuro ognuno avrebbe avuto il proprio robot personale così come adesso abbiamo un personal computer: essi sarebbero stati in grado di adattarsi su misura per aiutarci in qualunque tipo di attività. Le parti costituenti di questi robot sarebbero state abbastanza piccole da riassemblarsi a piacimento per assumere la struttura necessaria al compito da svolgere. Questo, tuttavia, a patto che venissero risolte alcune questioni: una maggiore comprensione dell'interazione tra computazione e mondo fisico era tra queste. In particolare, la riproduzione della coordinazione occhio-mano, che nell'essere umano è fondamentale, veniva indicata come una delle sfide più difficili.

Rus elencava alcuni dei successi ottenuti dalla robotica (“things we take for granted”³²⁵) e senza i quali sarebbe stato impossibile sperare in ulteriori progressi: i GPS, il WiFi, i processori GHz, il Gumstix, ecc. Questi, sviluppati nel corso del decennio precedente, si erano rivelati un'autentica manna per la ricerca.³²⁶ Dai loro successi erano nate la capacità di localizzazione, di pianificazione e di navigazione, oltre che la percezione

³²⁵ Houston 2006c.

³²⁶ Ibid.

3D. Ma l'ostacolo più grande era quello connesso all'incertezza: il mondo è imprevedibile e soggetto a continui mutamenti, ed era compito della robotica costruire robot in grado di gestirlo.

Un'altra sfida era costituita dalla percezione. Sebbene la robotica avesse compiuto grossi passi in avanti in campi quali la navigazione, altri campi altrettanto importanti, come quello della manipolazione, erano ancora molto indietro. Questo problema sarebbe stato impossibile da aggirare senza prima avere sviluppato algoritmi in grado di migliorare l'interazione tra uomo e macchine, e senza una migliore conoscenza delle connessioni tra percezione e azione. Rus dimostrava una certa sfiducia a riguardo:

“Simulations are always doomed to succeed”.³²⁷ Buoni risultati ottenuti in laboratorio non portavano quasi mai a risultati soddisfacenti nel mondo reale, nell'applicazione pratica.

Il problema di molti robot, soprattutto nei decenni passati, risiedeva nel fatto di non essere stati costruiti con l'obiettivo esplicito di comprendere l'intelligenza umana. Gli automi del secolo scorso precedentemente citati ne erano un esempio lampante, con la loro ripetizione meccanica di movimenti specifici per il raggiungimento di obiettivi altrettanto specifici. Le domande fondamentali nella progettazione dei robot avrebbero dovuto essere del tipo: qual è l'ambiente in cui interagisce un robot? E' completamente noto? In che modo può interagire con esso?

Durante la conferenza, Rus mostrò alcuni suoi lavori con i robot auto-

³²⁷ Ibid.

organizzati. L'idea di base era che le singole parti, per essere definite davvero intelligenti, avrebbero dovuto sapere come riorganizzare l'intera struttura. L'esempio usato era quello del robot che, arrivato sulla riva di un fiume, sarebbe stato in grado di costruire da solo un ponte.

Rus concludeva il suo intervento con alcune previsioni per il futuro. La robotica sarebbe stata in grado di fornire nuove teorie per l'intelligenza, se avesse saputo tener conto della situazionalità e dell'incorporamento del mondo reale. Uno sviluppo siffatto avrebbe permesso inoltre il superamento del concetto di intelligenza così com'era stato definito da Turing.

Un utilizzo più pervasivo di robot così concepiti avrebbe portato notevoli miglioramenti anche a livello ecologico. Secondo Rus, i biologi erano da sempre eccessivamente dediti a operazioni semplici come la raccolta di dati, mentre il loro lavoro non avrebbe dovuto limitarsi al semplice monitoraggio. L'utilizzo dei robot avrebbe permesso di riparare i processi danneggiati del sistema.

I robot avrebbero, infine, migliorato e prolungheranno la vita umana sotto tutti gli aspetti: la convergenza tra DNA e silicio era solo questione di tempo.

L'*AI@50* non fu solo una grande glorificazione dei successi raggiunti in quei cinquant'anni: rappresentò anche un perfetto manifesto di tutti quegli aspetti controversi, privi di posizione univoca, che caratterizzavano l'AI fin dagli esordi. Le diverse aree di ricerca al suo interno spesso non

collaboravano (e continuano tuttora a non collaborare) tra loro; i ricercatori utilizzavano differenti metodologie pur perseguendo gli stessi scopi, e non vi era ancora (e mai vi è stata) una teoria dell'intelligenza o sull'apprendimento univoca che unisse la disciplina. Anche a distanza di mezzo secolo, da questo punto di vista la situazione non era cambiata molto rispetto all'immediato post-Dartmouth.

Una delle controversie, forse la più significativa, riguardava la base su cui fondare gli studi sull'AI: se sulla logica o sulla probabilità. Sebbene l'AI delle origini fosse legata all'elaborazione simbolica e quindi alla logica non-monotona, le si erano affiancati, nel corso degli anni, metodi alternativi di tipo quantitativo (*certainty factors*, reti bayesiane, funzioni di credenza), basati sull'approccio probabilistico.

McCarthy continuava, ovviamente, a essere un fervente sostenitore dell'approccio logico. Ronald Brachman, capo dei Yahoo! Labs, nel suo intervento dal titolo *A Large Part of Human Thought*, sostenne che l'idea centrale della proposta del 1956 fosse che gran parte del pensiero umano “consists of manipulating words according to rules of reasoning and rules of conjecture” e che questa idea si sarebbe rivelata fondamentale per tutte le ricerche future. Risiedeva in ciò la rivoluzione dell'AI o, per usare le parole di McCarthy, la sua contro-rivoluzione, in quanto attacco al comportamentismo, posizione dominante nella psicologia degli anni Cinquanta.

Gran parte del pensiero umano è legato al linguaggio, ma vi sono anche casi in cui questa coesistenza viene sostituita dal solo linguaggio. Brachman citava la Proposta di Dartmouth:

Forming a generalization consists of admitting a new word and some rules whereby sentences containing it imply and are implied by others.³²⁸

Da questa affermazione si possono ricavare due concetti fondamentali: quello di “penetrabilità cognitiva” (*cognitive penetrability*), secondo cui il linguaggio può produrre un gran numero di comportamenti umani; e quello di “conoscenza delle cose” (*knowledge matters*). Durante la conferenza vennero usate come esempio due frasi grammaticalmente molto simili per esprimere meglio quest’ultimo concetto: “the city council refused to grant the protesters a permit because they feared violence” e “the city council refused to grant the protesters a permit because they advocated violence”. In entrambi i casi, chi legge intende il pronome “they” riferito a soggetti diversi, e questo perché il nostro background di conoscenze universali cambia l’interpretazione degli input linguistici, anche se questi, dal punto di vista linguistico, sono pressoché identici.

Buona parte del pensiero umano consisterebbe quindi in un “deposito” (*store*) di conoscenze rappresentate, e le procedure operate su di esse creano nuove conoscenze (inferenze) che influenzano le nostre decisioni sulle azioni da compiere. Il punto è capire cosa possiamo conoscere, e come, e

³²⁸ McCarthy et. al. 1955.

cosa ne consegue a partire da ciò che è noto: furono proprio la rappresentazione della conoscenza e il suo operare l'oggetto di studio principale della prima AI. E' dall'idea più elementare, formata da un semplice sillogismo come "Socrate è un uomo e tutti gli uomini sono mortali; quindi Socrate è mortale", che si arriva a immaginare altre possibilità, a una nuova visione del mondo. E da questa nuova idea nascono la scrittura, la lettura, la cultura e la scienza. Avere rappresentazioni della conoscenza disconnesse dalla realtà empirica implica l'intelligenza, la cultura e la società. E' possibile usare le strutture della conoscenza per ottenere inferenze proiettate anche molto in là nel futuro, per giungere mentalmente a delle conclusioni e prendere decisioni di conseguenza. Questa, secondo Brachman, era una delle idee più originali della conferenza del 1956, e aveva avuto una profonda influenza nella creazione dei Database management systems, degli *expert systems*, dell'OWL (Ontology Web Language), della robotica cognitiva ecc.

Percezione e linguaggio creano "memorie" utilizzabili, utili per compiere decisioni nel momento presente, o per pianificare azioni future in astratto. Si tratta di trarre vantaggio dalle azioni passate o, per meglio dire, dall'esperienza e dall'apprendimento che da essa ne derivano. L'esplicita rappresentazione formale di ciò che una macchina conosce e crede è di fondamentale importanza nella creazione di una macchina che può fare tutto ciò che fanno gli uomini.

Riassumendo: le nostre azioni sono condizionate da ciò che conosciamo e crediamo; almeno una parte di questa conoscenza è data in forma linguistica; questo ci permette di focalizzarci su cosa deve essere necessariamente conosciuto, e ciò dà origine a un approccio “basato sulla conoscenza” (*knowledge-based*), lo studio dell’intelligenza a partire da come e da quello che conosciamo. Secondo Brachman, quest’idea caratterizzò fin da subito l’AI, a differenza di altri campi come le neuroscienze computazionali, basati su modelli biologici.

David Mumford, professore emerito di matematica applicata, nel suo intervento dal titolo *What Is The Right Model for “Thought?”*, sosteneva di contro che gli ultimi 50 anni avevano comportato la graduale sostituzione della logica debole con il metodo probabilistico.

Tutto partiva da una domanda fondamentale: il pensiero è basato su regole logiche o su inferenze statistiche? Si tratta sicuramente di modelli molto diversi e, per Mumford, probabilmente incompatibili. La logica ha molte varianti: classica, modale, temporale, ecc. I modelli stocastici sono invece strutturati in maniera diversa. La deduzione logica è fragile e spietata, mentre invece l’inferenza statistica è più flessibile e adattiva:

Logic/rule-based thinking is fun, but is it real? To grammarians, rules are the deep and very real way to understand natural language grammars, but language is a maze of exceptions within exceptions.³²⁹

³²⁹ Ibid.

Assumendo questo punto di vista, le persone ragionerebbero più seguendo la logica bayesiana, per cui si creano delle inferenze sulla base di ipotesi sul contesto. Nel caso, per esempio, in cui si senta un suono in cui manca una consonante “The ?eel is on the shoe”, il suono percepito sarà “The heel is on the shoe”. Le persone sottoposte a un test del genere non diranno di aver pensato che la parola dalla consonante mancante sia “wheel” o “steel”: sanno che le scarpe hanno il tacco (*heel*), e così assumono, in assenza di altri dati, che la consonante mancante sia la “h”.

Sull'utilizzo dell'inferenza statistica, Mumford sosteneva che in qualunque situazione “reale” vi sono sempre una gran quantità di dati ridondanti (per non dire depistanti), e tantissime variabili in gioco. In queste situazioni, tuttavia, la statistica non sarebbe di nessun aiuto. Bisognerebbe, invece, fare una campionatura dell'intera distribuzione probabilistica e operare un filtraggio particellare (*particle filtering*) per ottenere un risultato. E' più o meno così che opera il cervello, e ciò gli è possibile perché non compie solo un'operazione per volta. Non vi è nessuna “grandmother cell” che si attiva quando si pensa alla propria nonna: il cervello gestisce centinaia di variabili contemporaneamente, filtrandole.

In definitiva, sebbene Mumford parteggiasse per il metodo probabilistico, la scelta su quali dei due metodi utilizzare, vista la scarsa conoscenza del cervello umano, era posticipata.

Eugene Charniak, nel suo intervento *Why Natural Language Processing is*

Now Statistical Natural Language Processing, sostenne questa posizione spiegando come l'elaborazione del linguaggio naturale (*Natural Language Processing*) avesse subito una rivoluzione nei 15 anni precedenti, passando da un dominio basato sulla logica a uno basato sulla statistica (*statistical natural language processing*).

Il *parsing* (o analisi sintattica) è un processo che analizza un flusso continuo di dati in ingresso (file, pagine web o input da tastiera) in modo da determinare la sua struttura grazie a una grammatica formale di riferimento. Il processo di *parsing* consiste nel passare dalla semplice stringa di parola al suo albero sintattico (*parse tree*), ovvero la rappresentazione della sua struttura sintattica. La frase verrà scomposta in sostantivi, verbi, ecc., in una completa analisi della sua struttura ad albero. Il *parsing* permette quindi di farsi una prima idea circa il significato della frase, scomponendo il significato generale nel significato delle singole parti. Ma senza conoscere le regole del linguaggio e della grammatica, sarà impossibile risalire a un'interpretazione corretta. Man mano che si aggiunge complessità, una singola frase può avere diverse interpretazioni: in questi casi, analisi diverse produrranno significati diversi. Data, ad esempio, la frase "The salesman sold the dog biscuits", potrebbero nascere dei dubbi come: il negoziante vende cibo per cani? Ha venduto cibo a un cane? Ha venduto un cane di nome Biscuits?

Il *parsing* pre-2006 assumeva che vi fossero solo poche interpretazioni

corrette per ogni frase, e gran parte della ricerca si concentrava sulla ricerca di algoritmi utili a questo scopo. Ma fino agli anni Novanta del secolo scorso non vi era ancora nessun *parser* (cioè nessun programma che esegue l'analisi sintattica della frase) in grado di trovare un'interpretazione semi-ragionevole per ogni frase contenuta, ad esempio, sulla prima pagina del *New York Times*, cosa che Charniak definiva “a perfectly reasonable requirement”.³³⁰

Sul perché fosse così difficile creare un programma del genere, Charniak rispose che, prima di tutto, le grammatiche “perdono” (*leak*), come un recipiente d'acqua pieno di buchi: è difficile concepire una grammatica in grado di gestire tutta la complessità del linguaggio naturale. Il secondo problema è quello dell'ambiguità. Per evitare che una grammatica “perda” bisognerebbe creare un *parser* esageratamente complesso: un compito quasi impossibile.

La conclusione cui giungeva Charniak era che, per trovare l'interpretazione più probabile per una stringa di input, bisognava riformulare il problema in chiave statistico-probabilistica. Il contesto grammaticale probabilistico, privo di riferimenti al contesto (*context-free*), offriva un insieme di regole che avrebbero definito le derivazioni per le stringhe formulate in linguaggio grammaticale, con tali derivazioni rappresentate dagli alberi: a quel punto sarebbe stata la probabilità, e non la logica, a essere interpretata.

Nel suo discorso, Charniak faceva riferimento al Penn Treebank, l'enorme

³³⁰ Houston 2006f.

database di alberi sintattici contenente più di un milione di parole e compilato interamente da esseri umani. Il materiale di riferimento usato dal Penn Treebank era in gran parte costituito da riviste, e gli ultimi *parsing* statistico-lessicali effettuati sul *Wall Street Journal* avevano raggiunto una certa precisione, facendo ben sperare per il futuro.

Ma ancora, alla domanda “come cogliere davvero il significato di una frase?”, Charniak rispondeva che bisognava procedere per gradi, partendo da ciò che si conosce. Ad esempio, se una frase contiene la parola “domani”, si può già capire il suo significato temporale.

Un altro punto di disaccordo, collegato alla questione della logica contro la probabilità, era quello che vedeva contrapposti psicologia e paradigma pragmatico. Pat Langley, Consulting Professor of Symbolic Systems presso la Stanford University, nel suo intervento *Intelligent Behavior in Humans and Machines*, affermava il bisogno di porre nuova enfasi sulla psicologia cognitiva. Allen Newell e Herbert Simon, tra i primi pionieri dell'AI, vedevano se stessi come psicologi, e in quanto tali provarono a spiegare il pensiero umano. Si trattava, a quel tempo, di un approccio tenuto in grande considerazione, e sicuramente quello di Newell e Simon non fu un caso isolato, visto che gran parte dei primi studi sulla rappresentazione (e il modo di replicarla meccanicamente) si occuparono spesso delle strutture e dell'organizzazione della conoscenza umana. Anche se è vero che non tutti gli studiosi dell'AI ebbero (o avrebbero avuto) questo tipo di

preoccupazioni psicologiche, è innegabile la loro influenza in tutte le ricerche successive.

Langley argomentava che questi primi studi, incentrati sulla capacità umana di *problem-solving*, avevano portato a intuizioni fondamentali nel campo della ricerca euristica, per quanto riguardava i rapporti tra stato-spazio (*state-space*) e scopo-direzione (*goal-directed*). Nel corso degli anni Ottanta del secolo scorso vi erano stati molti progressi in quella direzione, con lo sviluppo della capacità di elaborazione delle macchine a partire da certe conoscenze prendendo in prestito alcune idee e concetti provenienti dalla psicologia: ne sono un esempio gli *expert systems*, il ragionamento analogico, quello basato su modelli e gli studi sul linguaggio naturale.

La prima AI fu modellata sull'apprendimento umano, sulla capacità di *problem solving*, sul linguaggio naturale e sulla scoperta scientifica.

Secondo Langley, l'AI moderna si era invece allontanata dalla cognizione umana, ignorando i progressi della psicologia, nonostante il pesante debito storico. Ne era un esempio il problema della rappresentazione della conoscenza, dove l'attenzione si rivolgeva ora al modo di rendere l'elaborazione efficiente, o allo stesso *problem solving* e alla pianificazione, dove ci si concentrava maggiormente sul metodo statistico, che non ha nessun collegamento con le strutture linguistiche o psico-linguistiche. Nell'apprendimento meccanico le tecniche statistiche erano diventate dominanti, ma avevano l'inconveniente di richiedere un numero maggiore

di dati e di esempi-guida rispetto a quelli richiesti agli esseri umani per raggiungere lo stesso risultato.

Per Langley, questo spostamento di paradigma era da imputare agli sviluppi della tecnologia. Computer più veloci possono analizzare un gran numero di dati in poco tempo, permettendoci di elaborare più modelli di predizione.

Ma tale potenza di calcolo andrebbe combinata con intuizioni propriamente psicologiche. In più, l'AI veniva (e viene tuttora) studiata e insegnata quasi esclusivamente nelle facoltà d'informatica, che spesso sono una semplice estensione di quelle di matematica. I matematici, per Langley, hanno un atteggiamento mentale tendente all'arrendevolezza, nel loro ridurre il lavoro a un piccolo numero di classi di problemi, e sono generalmente scettici nei confronti della psicologia. Nonostante i successi commerciali dell'AI, ciò tende a incoraggiare i lavori minori, mentre in ambito accademico si dovrebbe avere il coraggio di voler raggiungere dei risultati più ambiziosi.

Uno dei motivi per rinnovare l'interscambio tra AI e psicologia era senza dubbio quello di capire meglio la cognizione umana. Le abilità umane sono una fonte inesauribile di sfide per la ricerca, e mostrano come le persone possano compiere imprese straordinarie.³³¹ Dall'altra parte, la psicologia cognitiva ha anche un proprio diritto di rappresentanza, avendo dimostrato come gli esseri umani ricorrono all'euristica (e non agli algoritmi) durante il *problem solving*. Nella teoria dell'apprendimento, la psicologia cognitiva ha mostrato come gli individui formino nuove strutture cognitive durante il

³³¹ Ibid.

processo di *problem solving*, e come l'apprendimento sia incrementato e intervallato dalle *performance*, ovvero dalla risoluzione di quegli stessi problemi

L'AI avrebbe dovuto, insomma, tornare alle proprie radici. Concludeva Langley: "The AI curriculum should include cognitive psychology, structural linguistics, logical reasoning, and the philosophy of mind".³³²

Altri ricercatori, tuttavia, erano più propensi a esplorare cosa sarebbe successo nel caso in cui si fosse provato a raggiungere un simile obiettivo percorrendo vie che si potrebbero definire "non umane": Peter Norvig, Director of Research di Google, era tra questi.

Nel suo intervento *Web Search as a Product of and Catalyst for AI*, egli suggeriva come la ricerca (*searching*, cioè la ricerca d'informazioni), in particolare vista la grande reperibilità di dati sul web, mostrasse segnali incoraggianti sulla possibilità di risolvere i problemi tradizionali dell'AI senza ricorrere alla psicologia. L'apprendimento meccanico e l'elaborazione del linguaggio naturale erano diventate le tecnologie centrali della ricerca. Google, il cui scopo è quello organizzare tutte le informazioni possibili per renderle universalmente accessibili e fruibili, offre ad esempio la funzione "elementi previsti" (*predicted items*), che porta immediatamente a una classificazione dell'oggetto cercato: cercando "Armani" su Google, si ottengono come suggerimenti altre marche come Gucci, Prada, ecc.

Norvig era promotore dell'approccio basato sul corpus (*corpus-based*) per

³³² Ibid.

la comprensione del linguaggio, in opposizione a quello basato sulla conoscenza. Un corpus è una grande collezione di testi, come lo sono le miliardi di pagine che compongono il World Wide Web. Il testo è scritto da umani per gli umani, e il compito del software è quello di rendere facilmente reperibile l'informazione cercata. Questo approccio richiede l'uso della statistica e dell'apprendimento per sfruttare il corpus, e poggia su modelli probabilistici del linguaggio che possono essere appresi dai dati.³³³

L'uso di questo approccio ha permesso di raggiungere risultati sorprendenti. Uno di essi è la correzione ortografica della parola cercata tramite il motore di ricerca, nel caso di una digitalizzazione sbagliata. In questo caso, Google costruisce un modello probabilistico basato sul gran numero di variazioni ortografiche contenute nel corpus. Il vantaggio dell'approccio basato sul corpus è quello di non cercare i nomi su un dizionario, ma in un contesto, in riferimento alle parole vicine.

Un'altra funzione di Google è quella che permette di estrapolare fatti a partire da semplici testi: nel caso in cui, ad esempio, si voglia sapere a quanto ammonta la popolazione giapponese, basterà cercare "what is population of Japan" e Google troverà l'informazione inserendola all'inizio dei risultati di ricerca.

E' possibile anche tradurre con un livello abbastanza alto di precisione, per

³³³ Russell, Norvig 2005: 547.

esempio, dall'arabo all'inglese, utilizzando i soli metodi statistici anche se nessuno nello staff di Google parla l'arabo. Ciò è possibile tramite la raccolta di testi paralleli (ovvero lo stesso testo esistente in due lingue diverse, come può esserlo un quotidiano) e produrre un algoritmo probabilistico.

L'approccio basato sul corpus differisce sostanzialmente da quello basato sulla conoscenza anche per via di aspetti più pragmatici ed economici. Certi concetti sono del tutto assenti nei libri usati generalmente come riferimento. Il concetto che "l'acqua scorre verso il basso", ad esempio, non è qualcosa che si può cercare in un'enciclopedia. In più, anche nel caso ci si volesse servire della conoscenza esistente (cioè di testi "ufficiali"), come può esserlo un libro di chimica, costerebbe circa 10.000 dollari codificare una singola pagina. Il problema risiederebbe quindi nel modo di ottenere informazioni e conoscenze cercando di contenere i costi.

Google ha il vantaggio di possedere un corpus d'informazioni enorme, molto più grande di quello che può essere contenuto in un'enciclopedia o un libro di testo. Tanto è più grande il corpus, tanto maggiore è la possibilità di evitare risultati di ricerca ambigui.

Norvig fece riferimento al libro di Andy Clark *Being There: Putting Brain, Body, and World Together Again*. Gli esseri umani non sono intelligenti come credono di essere: hanno semplicemente capito alcune cose intelligenti che si possono fare con il linguaggio e la matematica, ma ci

sono sempre dei “prolungamenti”, dei sistemi esterni, che amplificano la loro intelligenza. In questo, siamo del tutto giustificati (e anzi incentivati) a creare macchine intelligenti ricorrendo a metodi alternativi e poco “naturalisti”.

Norvig chiuse ironicamente l’intervento con un’altra citazione, stavolta di Sergey Brin, uno dei fondatori di Google: “In the future search engines should be as useful as HAL in 2001, but hopefully they won't kill people”.³³⁴

Infine, nel 2006 si discusse anche su come le reti neurali avrebbero potuto rivelarsi utili per il raggiungimento della perfetta AI. L’argomento fu trattato da Simon Osindero e Geoffrey Hinton nell’intervento *From Pandemonium to Graphical Models and Back Again*. Osindero iniziò ponendo subito una domanda: “How to build a perceptual system , a pandemonium architecture, after Selfridge?”.³³⁵

Le vecchie reti neurali erano costituite da *input layer*, *hidden layer* e *output layer*, oltre che da un algoritmo di *backpropagation* che fungeva da surrogato per l’apprendimento. Bisognava chiedersi cosa non andasse in questo modello.

Una rete neurale necessitava di dati di allenamento (*training corpus*)

³³⁴ Il riferimento è al film *2001: Odissea nello spazio*, in cui il supercomputer HAL (acronimo di *Heuristic ALgorithmic*) tenta di eliminare l’intero equipaggio dell’astronave Discovery.

³³⁵ Houston 2006g.

“etichettati”³³⁶ (*labeled*), divisi per categorie. Ma ciò presentava degli inconvenienti. Prima di tutto, questo sistema era inefficace, perché le etichette di per sé non fornivano abbastanza informazioni sui dati da elaborare. Seconda cosa, il tempo di apprendimento (*learning time*) era eccessivamente sbilanciato quando si aveva a che fare con più di 2 o 3 layer. Terza cosa, i neuroni artificiali avevano bisogno di trasmettere il segnale al neurone successivo e l’informazione retro-propagata a quelli precedenti, mentre i veri neuroni non si comportano così.

Osindero e Hinton ipotizzarono che una possibile soluzione avrebbe potuto venire da una versione ristretta della macchina di Boltzmann: una rete neurale in cui i neuroni aggiornano il proprio stato in base a regole stocastiche e non deterministiche.³³⁷

Una rete del genere così concepita avrebbe contato solo due layer: uno preposto alle unità nascoste (non connesse tra di loro); e uno per le unità visibili. Iniziando dai dati di allenamento del layer visibile (ovvero le immagini della realtà), queste sarebbero state caricate ed elaborate dal layer nascosto, che a sua volta avrebbe aggiornato in tempo reale il layer visibile per ottenere una “ricostruzione” della realtà, ripetendo il ciclo all’infinito.

Una rete neurale così progettata avrebbe potuto imparare a riconoscere le caratteristiche dei dati di allenamento, cercando di “capire” tutte le

³³⁶ Prima di entrare in funzione, una rete neurale viene solitamente sottoposta a una fase di allenamento (*training*) durante la quale vengono testate le correlazioni esistenti tra i segnali di input e di output.

³³⁷ Spezzano, Talia 1999: 144.

percezioni in base al proprio *corpus* di dati: se, ad esempio, la rete avesse posseduto la nozione del numero 2, e le venisse dato da elaborare il numero 3, avrebbe potuto riconoscere alcune caratteristiche di quest'ultimo numero a partire da certi tratti comuni che lo legano al primo.

Sarebbe stato possibile anche costruire modelli più complessi dotati di layer nascosti multipli. Questo approccio sembrava più promettente rispetto a quello basato sulla retro-propagazione, perché vi era un solo tipo di segnale trasmesso attraverso il sistema. Una macchina del genere avrebbe potuto “fantasticare” (*fantasize*) output capaci di ricostruire la realtà. Per esempio, se a questo sistema venisse dato il simbolo 2, essa potrebbe creare un output a forma di 2. Allo stesso modo, se venisse dato al sistema un 2 scritto a mano, essa indovinerebbe la maggior parte delle volte che quel simbolo rappresenta un 2. Ciò non comporterebbe, tuttavia, che il sistema possieda o esprima alcuna comprensione sul concetto di “duezza” (*two-ness*).³³⁸

Sempre riguardo le reti neurali, sia Terry Sejnowski sia Rick Granger spiegarono che nell'ultimo decennio si era giunti a una migliore comprensione del cervello umano, e che queste nuove conoscenze si sarebbero rivelate fondamentali per la creazione di modelli computerizzati di attività intelligenti.

L'intento dichiarato da Sejnowski durante la conferenza era proprio quello di capire meglio il funzionamento del cervello. Il problema risiedeva nel

³³⁸ Houston 2006g.

conoscere quali fossero i giusti principi sottesi al suo funzionamento, e come capirli. Secondo il suo punto di vista, solo una comprensione piena di quei principi avrebbe permesso la loro “ricostruzione”, ovvero la possibilità di poterli replicare meccanicamente.

Sejnowski disse di aver chiesto una volta ad Allen Newell perché avesse deciso di voltare le spalle alle neuroscienze. Newell rispose che a quel tempo (ovvero agli albori dell'AI) il campo di ricerca era ancora troppo giovane per poter fornire delle informazioni utili. Nel 2006 tutto sembrava cambiato, perché si disponeva di migliori tecniche per analizzare il cervello a qualunque livello di indagine, da quello molecolare fino alle strutture complesse.

Nel libro che dava il titolo all'intervento, *A Critique of Pure Vision*, Sejnowski, insieme a Patricia Churchland e Vilayanur Ramachandran, sosteneva che vi sia un errore di fondo nell'assumere la vista come un'architettura *feedforward*. Fra i vari tipi di architetture, quella del *feedforward* (o a propagazione in avanti) è una delle più frequentemente usate nelle ricerche sulle reti neurali. Essa è formata da un insieme (o strato) di unità di input, da uno strato di unità di output e da uno o più strati intermedi di unità interne o nascoste. Ogni unità di ciascuno strato è connessa con tutte le unità dello strato successivo, a partire dallo strato di input fino a quello di output. In ogni ciclo, l'attivazione si propaga in avanti dalle unità di input ai successivi strati di unità interne fino a raggiungere le

unità di output.³³⁹ La vista, tuttavia, non rientrerebbe in questo schema, e il fenomeno della cecità al cambiamento (*change blindness*) confuterebbe quest'argomento. Se, ad esempio, si proietta su uno schermo la stessa immagine in alternanza con una identica, salvo che per un piccolo particolare, gli osservatori, in genere, non inquadrano immediatamente l'oggetto mancante, per il semplice fatto che l'immagine è in genere così piena di particolari che risulta impossibile memorizzarli tutti la prima volta. Ora sappiamo quali sono i meccanismi del cervello che permettono a questo fenomeno di verificarsi: si tratta di un fenomeno neurologico, ed è l'attenzione che modula la frequenza di scarica dei neuroni nell'apparato di riconoscimento. Si tratta quindi di un'attività dovuta a un segnale interno, non derivante dal mondo esterno.

Sejnowski citò anche gli esperimenti condotti sugli animali riguardanti lo studio dell'apprendimento per differenze temporali (*temporal difference learning*). Questi si sarebbero potuti rivelare fondamentali per lo sviluppo dell'apprendimento meccanico, nella ricerca degli algoritmi adatti a simulare il comportamento umano: le tecniche skinneriane non erano inefficienti in sé, ma andavano riformulate in vista delle conoscenze acquisite negli anni.

Si trattava, comunque, di ipotesi per il futuro: nel 1956 così come nel 2006, a essere nota era solo la parte finale dei processi cognitivi, il risultato, quello che si “crea” dal cervello umano. Tutto il “lavoro pesante”, i processi

³³⁹ Parisi 2007.

fondamentali, avvengono sotto la sua superficie, che va scalfita per scoprire quali algoritmi vi siano nascosti: quello dell'apprendimento per differenze temporali sarebbe solo uno di questi.

In una conferenza dedicata al Congresso di Dartmouth e a tutto ciò che ha comportato, un intervento come quello di Sejnowski (quasi un'apologia di tutti quei metodi che l'AI degli inizi aveva rigettato) appariva quantomeno curioso, e la cosa non passò inosservata, tanto che l'obiezione da parte di uno dei presenti fu: "there's not a lot of clarity about what's AI, what's cognitive science, what's neuroscience. What you're doing is not obviously AI".

Sejnowski rispose così:

My goal is to understand how the brain works. If all you have is the input and output, you're doomed. You need a way of understanding how the whole system is organized in between.³⁴⁰

Nell'intervento su *Essential Circuits of Cognition: The Brain's Basic Operations, Architecture, and Representations*, Granger affermava che l'obiettivo dell'AI è sempre stato duplice: da una parte, la spiegazione formale dei meccanismi sottesi all'intelligenza umana (e animale); dall'altra, la costruzione di artefatti intelligenti basati su tali meccanismi. Quest'ultimo obiettivo, di tipo ingegneristico, potrebbe trarre beneficio da quello precedente, più propriamente scientifico: i sistemi di riconoscimento

³⁴⁰ Houston 2006h.

facciale e gli operatori telefonici automatizzati potrebbero anche essere considerati come un'avanguardia tecnologica difficilmente superabile, ma non è così. Noi sappiamo che questi sistemi industriali possono (e devono) essere superati per il semplice fatto che non possono competere con l'essere umano, con la sua intelligenza. L'uomo era ed è l'idea-limite cui ogni ricerca sull'AI deve necessariamente tendere.

I sistemi biologici riescono a realizzare le loro capacità cognitive esclusivamente attraverso i meccanismi cerebrali, tramite il funzionamento fisiologico di circuiti anatomici. I circuiti cerebrali sono, appunto, circuiti, e in quanto tali possono essere compresi in termini computazionali. Il decennio passato ha visto un'esplosione di nuove conoscenze nel campo delle neuroscienze, rivelando dati fondamentali per una migliore comprensione della struttura e delle proprietà di questi circuiti. Ai fini dell'AI, queste informazioni possono essere organizzate in tre categorie:

- Operazioni di base: riguardante gli operatori elementari che effettuano le operazioni mentali fondamentali.
- Architettura: riguardante la struttura organizzativa di controllo in cui gli operatori sono inseriti.
- Rappresentazione: riguardante le meccaniche con cui i ricordi e la conoscenza strutturata vengono immagazzinati e recuperati all'occorrenza.³⁴¹

La miriade di compiti svolti dall'intelligenza (riconoscimento visivo e

³⁴¹ Granger 2006.

uditivo, la pianificazione, il linguaggio, e molti altri) sono, con poche eccezioni, vaghi e privi di specificazione formale. A differenza di quanto avviene nella progettazione ingegneristica, in cui una specificazione formale precede solitamente la soluzione, questi compiti vengono studiati usando un unico (e possibile) punto di riferimento: l'osservazione diretta dei sistemi intelligenti (umani e animali) che li eseguono, con procedure che ci sono sconosciute. Questo approccio può essere chiamato ingegneria inversa (*reverse engineering*), ovvero l'osservazione del comportamento nel tentativo di capire i suoi meccanismi di base, e in genere veniva rifiutato proprio per via del suo basso grado di specificazione.

Ma i numerosi campi correlati alle neuroscienze (anatomia, fisiologia, biochimica, farmacologia, ecc.) avevano generato nell'ultimo decennio una profusione senza precedenti di nuovi dati, guadagnandosi a buon diritto il soprannome "the decade of the brain". Questo vasto tesoro d'informazioni avrebbe permesso per la prima volta approcci di tipo *bottom-up* alla comprensione dell'intelligenza, in maniera tutt'altro che vaga, tramite la caratterizzazione formale dei comportamenti derivanti dalla struttura anatomica e la maggiore comprensione del funzionamento fisiologico di circuiti biologici nelle loro architetture di sistemi endogeni. Tale compito si poneva in maniera "bizzarra" rispetto a quello dell'AI "classica". Mentre l'AI cercava da sempre di identificare quei meccanismi che avrebbero potuto portare a un risultato funzionale predeterminato, lo scopo di Granger

era quello di identificare quali risultati sarebbero emersi da una serie prestabilita di meccanismi biologici.

Il discorso si chiuse con delle considerazioni simili a quelle di Sejnowski: esistono degli algoritmi che possono essere derivati dai meccanismi cerebrali, e la ricerca di tali algoritmi era compito (anche) degli studi sull'AI.

Questa dialettica, fatta di scontri e posizioni anche diametralmente opposte, poteva essere vista come un segno di salute. Come disse Nils Nilsson, il summit rappresentava perfettamente i diversi percorsi che l'AI stava percorrendo, anche se probabilmente non tutti si sarebbero rivelati fruttuosi nel lungo termine. Ma dal momento che non si poteva determinare a priori quale sarebbe stato il migliore, avere tante possibilità rappresentava un vantaggio.³⁴²

Nonostante i diverbi e i contrasti, l'*AI@50* si configurò come erede spirituale del congresso del 1956 anche per un altro motivo: la visione comune che i computer potessero compiere compiti intelligenti. Forse, come affermò Moor, era proprio questa visione l'unica cosa a tenere unito il campo di ricerca, nel 2006 così come nel 1956.

³⁴² Houston 2006e.

3.3 Previsioni per il 2056

Durante l'*AI@50* vennero fatte anche molte predizioni sul futuro dell'AI. Quando fu chiesto loro come sarebbe stata l'AI tra 50 anni, i partecipanti al congresso originale mostrarono posizioni diverse. Secondo McCarthy, un'AI a livello umano sarebbe stata auspicabile per il 2056, a patto che si risolvessero i problemi esposti durante il suo intervento.

Selfridge dichiarò che in futuro le macchine sarebbero state capaci di una migliore pianificazione, incorporando qualcosa di simile (ma non paragonabile) alle sensazioni e agli affetti umani. Tutto ruotava intorno al concetto di apprendimento. L'essere umano è in grado di apprendere molte cose contemporaneamente, parallelamente. Ad esempio, un neonato nella culla muove il braccio e focalizza la visuale su un oggetto scintillante poco lontano, in modo da raggiungerlo. La prima volta, per errore, lo fa cadere a terra. Ma nel giro di una settimana riuscirà a raggiungerlo con più sicurezza. Saprà quanto e quando rallentare i suoi movimenti. E più avanti, crescendo, quando si ritroverà per la prima volta a calciare un pallone, ripeterà la stessa logica, rallentando la propria corsa un attimo prima di colpirlo, concentrandosi sull'oggetto ed evitando una caduta rovinosa. Utilizza, cioè, una tecnica utile, ma appresa in un altro contesto. Selfridge non si preoccupava di come il bambino riuscisse a far ciò, formalmente parlando: semplicemente, non era importante.

E' più importante notare che un apprendimento di quel tipo è impossibile da

applicare quando, per esempio, si mangiano gli spaghetti, perché l'apprendimento differisce per complessità. In questo senso, la lettura non è altro che il riconoscimento di figure: non si “legge” nel vero senso della parola. Guardiamo una parola o una frase, e la colleghiamo al suo significato, a un insegnamento o un apprendimento passato legato a quel determinato concetto.

Sul perché si apprende, la risposta di Selfridge fu: “to use technical terms, why do we give a shit?”³⁴³. Scopi e obiettivi giocano un ruolo fondamentale nell'apprendimento. Impariamo perché vogliamo qualcosa: “If you don't learn, it's probably because you don't care”.³⁴⁴

Non si ha mai un solo scopo, ma una gerarchia di scopi. Ogni azione o pensiero è la risposta a uno scopo, ovvero a qualcosa che si sta cercando di fare. Quando si impara, ad esempio, a giocare a scacchi, la vittoria è l'obiettivo principale, ma si dispiegano anche tutta una serie di sub-strutture per dei sub-scopi (come le singole mosse allo scopo, per esempio, di “mangiare” un pezzo importante) che eventualmente porteranno alla vittoria. Quindi, quando si prova ad apprendere qualcosa, lungo la via per il raggiungimento dello scopo si imparano tutta un'altra serie di cose necessarie per il raggiungimento del macro-scopo, e tutti questi apprendimenti minori sono fondamentali nel comportamento reattivo e ricco che caratterizza gli esseri umani.

³⁴³ Houston 2006a.

³⁴⁴ Ibid.

Per Selfridge, l'importanza degli scopi non andava sottovalutata.

Dall'apprezzamento della bellezza, scendendo in basso fino al movimento dei muscoli, tutto può essere spiegato come un circolo controllato di scopi: "Purposes get changed; circumstances arise".³⁴⁵ Contesto e scopo infatti si influenzano a vicenda: la nostra concezione di bellezza a 8 anni non è la stessa di quella che si forma a 18. Ogni scopo va collocato all'interno del proprio contesto, e ogni contesto va collocato intorno a uno scopo.

Bisogna considerare il cambiamento quale fattore fondamentale della scoperta:

Be sensitive to the role of change. Life is always changing. And that change changes your point of view and your purposes. Part of learning has to be to understand that and to keep going anyhow. Marvin Minsky put it well when he said "the best is the enemy of the good." Meaning the best implies an optimization, which, when used in machine learning, implies that the world is standing still. Which is not a safe premise.³⁴⁶

Bisognava trovare il modo di esprimere in un linguaggio computerizzato non solo quello che il software avrebbe dovuto fare, ma anche quello che avrebbe dovuto provare a fare. Durante ogni fase della sua elaborazione si sarebbe dovuto dire al software di fare quello per cui era stato programmato e, contemporaneamente, di provare qualcosa di nuovo. Un'altra caratteristica fondamentale dell'apprendimento è che singolarmente non si

³⁴⁵ Ibid.

³⁴⁶ Ibid.

impara mai tanto quanto si apprende culturalmente. Era necessario creare un linguaggio in grado di esprimere molti concetti, non solo quelli strettamente necessari:

Tell an AI program that it doesn't just want to do something, but to create a satisfaction parameter that will make it happier that it has done what it's tried to do. And remember, it's not just trying one thing, it's trying many things.³⁴⁷

Sempre per il futuro dell'AI, per Minsky era fondamentale la presenza di ricercatori brillanti che inseguissero unicamente le proprie idee, senza replicare quelle dei loro predecessori. Ricercatori del genere erano diventati merce rara, e la categoria sembrava più attratta dagli aspetti puramente accademici ed economici della ricerca. Dal suo punto di vista, l'AI non sarebbe mai stata una scienza fino a quando non si fosse preoccupata di approfondire sia quello che funzionava sia quello che non funzionava. Solomonoff credeva che la creazione di macchine davvero intelligenti non fosse troppo lontana o fuori portata. Il pericolo era più di natura politica: le tecnologie erano ormai in grado di dare un grande potere, misurabile, a individui e governi.

Ray Kurzweil offriva una visione più ottimista sul progresso e dichiarava che si sarebbe avuto un'AI capace di superare il test di Turing nel giro di 25 anni – una previsione su cui molti erano in disaccordo.³⁴⁸ Prevedere eventi

³⁴⁷ Ibid.

³⁴⁸ Nel 2014, Vladimir Veselov e Eugene Demchenko crearono il cleverbot Eugene Goostman, il quale riuscì a superare il test di Turing, convincendo il 33% dei giudici. Il caso fu però oggetto di aspre contestazioni, visto che Goostman simulava un ragazzo di 13 anni non madrelingua inglese,

futuri, in fondo, è sempre un azzardo. Simon predisse nel 1960 che un calcolatore sarebbe stato campione di scacchi nel giro di 10 anni. La previsione si rivelò giusta, ma non la data: avvenne solo 40 anni dopo. Per Kurzweil il ritmo della tecnologia era e sarebbe rimasto incostante, e i nuovi paradigmi si succedevano con una frequenza doppia rispetto al decennio precedente. Il ritmo stava accelerando, sia nello sviluppo tecnologico che nel suo impiego. Nella sua visione, quando il paradigma vigente aveva detto tutto, si era pronti per il prossimo, e il ciclo ripartiva. La legge di Moore rappresentava il quinto paradigma di cui si era a conoscenza, per quanto riguarda la tecnologia. Prima c'erano state le valvole termoioniche (o tubi a vuoto), andate via via in disuso man mano che la loro potenza aumentava. Poi fu la volta dei chip, e presto qualunque altro progresso sarebbe stato impossibile senza ricorrere a dei chip a tre dimensioni.

Tutti quei limiti che nel 2006 sembravano insuperabili (come la capacità di elaborazione dei computer) in futuro non sarebbero più stati così vincolanti.

A riguardo, Kurzweil si lanciava in altre predizioni: entro il 2020 si sarebbero potute simulare tutte le caratteristiche del cervello umano, grazie ai software di *scanning* che nel corso degli anni erano diventati sempre più avanzati, ottenendo anche qualche successo.

Sherry Turkle, professoressa di Social Studies of Science and Technology

“giustificando” così il fatto che non potesse (o sapesse) rispondere adeguatamente a determinate domande. Per maggiori approfondimenti: <http://www.wired.it/attualita/tech/2014/06/09/computer-superato-test-turing/>

al MIT, poneva invece l'accento sull'elemento umano, troppo trascurato nello sviluppo tecnologico. Le macchine rappresentavano ormai una realtà con le quali ci si doveva necessariamente confrontare, anche a livello umano. Nel futuro, con il progredire di queste macchine, una questione fondamentale avrebbe potuto riguardare non tanto le capacità effettive dei computer, ma la nostra vulnerabilità nei confronti di AI particolarmente sofisticate.

Le tecnologie non sono mai solo strumenti impersonali, ma “oggetti evocativi”. Le entità computazionali (computer, robot, ecc.) sono diventate ormai realtà della vita di tutti i giorni, compagni con cui doversi necessariamente relazionare, provocandone l'antropomorfizzazione.

Ogni artefatto relazionale è un oggetto evocativo. Queste macchine si presentavano ormai come se avessero degli stati d'animo interni influenzati dalle interazioni con gli umani. Non colpivano solo per via della loro intelligenza, ma anche per la loro sensibilità. Nel nostro interagire con loro e nel riflettere sulle relazioni che instauriamo con loro, apprendiamo di più su noi stessi, sui nostri stati interni.

Turkle prendeva come esempio alcuni giocattoli elettronici molto in voga nel 2006, come il Furby e il My Real Baby³⁴⁹: quest'ultimo non fu un grosso successo presso i bambini, semmai fra gli anziani nelle case di riposo. Bambini e anziani rientrano nella popolazione più fragile. Negli anziani in particolare un'interazione robotica può in certi casi compensare

³⁴⁹ Quest'ultimo mai distribuito in Italia.

un'assenza totale di interazioni. Ma un robot non sarà mai veramente empatico. Tutto ciò andava considerato e soppesato, moralmente parlando, al fine di chiedersi come e in che misura questi oggetti avrebbero potuto offrire una compagnia, una compensazione positiva, ai disabili, agli anziani, ai bambini.

L'arte e la scienza degli artefatti relazionali coincidono con l'arte e la scienza della comprensione della psicologia e dell'umana vulnerabilità.

L'interazione con questi oggetti sposta l'attenzione dalla psicologia del progetto alla psicologia dell'impegno e della relazione oggettuale. Nel parlare di artefatti relazionali, il discorso della vitalità si muove dalla responsabilità degli oggetti (fisica o psicologica) alla relazione della persona con l'oggetto. Questi oggetti sono in grado di premere i nostri bottoni darwiniani (*Darwinian buttons*), dal momento che rispondiamo loro come se fossero senzienti e dotati di sentimento.

Tuttavia, nella relazione educativa, i bambini dovrebbero imparare a distinguere tra l'animale vivo e il robot che simula la vita, tra l'amore animale e l'amore meccanico. I bambini, già dalla più tenera età, sono ormai abituati ad allevare (*to nurture*), letteralmente, i robot: pensano che un loro malfunzionamento sia dovuto a una ferita, mostrano belligeranza quando non lavorano come dovrebbero, e li assumono come interlocutori nella conversazione. Si potrebbe dire che questo è un atteggiamento comune a tutti, non solo ai bambini: secondo Turkle, la differenza fondamentale

risiedeva nel fatto che le nuove generazioni erano le prime a entrare in una simile relazione affettiva già dalla più tenera età.

Gli artefatti relazionali si presentano come possibile oggetti-Sé. Essi costituiscono una relazione affettiva positiva quando l'oggetto è una persona, in quanto riflesso di un'esperienza narcisistica che proietta il sé sugli altri. Bisognava considerare cosa succede quando l'oggetto-sé cessa di essere una persona, sostituito da un artefatto; se vi fossero dei possibili sbocchi terapeutici (e quanto opportuni), pur nell'illusione, consapevole o meno, che tali artefatti fossero capaci di empatia. Era anche necessario considerare come una simile relazione avrebbe permesso lo sviluppo di nuovi modi di sentire, di empatizzare verso esseri diversi da quelli umani, sviluppando tipi di relazioni prima mai esistite nella storia dell'essere umano.

Bibliografia

- Aa. Vv., *Almanacco Letterario Bompiani*, Milano, Bompiani, 1962.
- Aa. Vv., *Atti del convegno sul tema: L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, Roma, Accademia Nazionale dei Lincei, 1968.
- Aa. Vv., *L'uomo e la macchina. Atti del XXI Congresso Nazionale di Filosofia, Pisa 22-25 aprile 1967*, Torino, Edizioni di «Filosofia», 1967 (3 voll).
- Aa. Vv., *La cultura informatica in Italia. Riflessioni e testimonianze sulle origini. 1950-1970*, Torino, Bollati Boringhieri - Fondazione Adriano Olivetti, 1993.
- Aa. Vv., *Vittorio Somenzi 1918-2003: antologia e testimonianze*, Mantova, Fondazione Banca agricola mantovana, 2011.
- F. Accame, “Silvio Ceccato”, in *Dizionario Biografico degli Italiani*, 2014: [http://www.treccani.it/enciclopedia/silvio-ceccato_\(Dizionario-Biografico\)/](http://www.treccani.it/enciclopedia/silvio-ceccato_(Dizionario-Biografico)/)
- E. Agazzi, “Simulazione del pensiero e intelligenza artificiale”, in Aa. Vv., *L'uomo e la macchina. Atti del XXI Congresso Nazionale di Filosofia, Pisa 22-25 aprile 1967*, vol. 2, Torino, Edizioni di «Filosofia», 1967, pp. 45-48.
- E. Agazzi, V. Fortunati, *Memoria e saperi: percorsi transdisciplinari*, Roma, Meltemi Editore, 2007, p. 490.
- J. A. Anderson, E. Rosenfeld, *Talking Nets: An Oral History of Neural Networks*, Cambridge, MIT Press, 1998, pp. 48-49.
- P. Antonello, “Cibernetica e fantasmi. Italo Calvino fra mito e numero”, in P. Antonello, *Il ménage a quattro. Scienza, filosofia, tecnica nelle letterature italiana del Novecento*, Firenze, Le Monnier, 2005, pp. 169-242.
- P. M. Asaro, “From Mechanisms of Adaptation to Intelligence Amplifiers: The Philosophy of W. Ross Ashby”, in P. Husbands, O. Holland, M. Wheeler, *The Mechanical Mind in History*, Londra, MIT Press, 2008, pp. 149-184.
- W. R. Ashby, “Design for an Intelligence Amplifier”, in C. Shannon, J. McCarthy, *Automata Studies*, Princeton University Press, 1956, pp. 215-234.
- W. Aspray, *An Interview with John McCarthy*, Palo Alto, Charles Babbage Institute, 1989, consultabile presso il sito: <http://conservancy.umn.edu/bitstream/handle/11299/107476/oh156jm.pdf?sequence=1&isAllowed=y>

W. Aspray, "The Scientific Conceptualization of Information: A Survey", in *IEEE Annals of the History of Computing*, vol. 7, n. 2, 1985, pp. 117-140.

A. Astin, L. Cohen, J.W. Forrester, A. W. Jacobson, J. Mauchly, R. E. Meagher, D. Sayre, "The Impact of Computers on Science and Society", in *IRE Transactions on Electronic Computers*, EC-5, 1956, pp. 142-158.

S. Augarten, *Bit by Bit*, New York, Ticknor & Fields, 1984.

N. Balestrini, "Tape Mark I", in Aa. Vv., *Almanacco Letterario Bompiani*, Milano, Bompiani, 1962, pp. 145-151.

Y. Bar-Hillel, "The Present State of Research on Mechanical Translation", in *American Documentation*, vol. 2, n. 4, 1951, pp.229-237.

J. Bernstein, *Uomini e macchine intelligenti*, Milano, Adelphi, 1990.

M. L. Bianca, "Artificial Intelligence", in M. Alai, M. Buzzoni, G. Tarozzi, *Science Between Truth and Ethical Responsibility. Evandro Agazzi in the Contemporary Scientific and Philosophical Debate*, Springer, 2015, pp. 91-104.

C. Bo, "Gli ingegneri del verso", in *L'Europeo*, 14 gennaio 1962.

G. Bontadini, "Il processo di banalizzazione della filosofia contemporanea", in *Rivista di filosofia neo-scolastica*, n. 47, 1955, pp. 291-301.

A. Borsellino, "Neuroni, percezioni, intelligenza", in Aa. Vv., *Atti del convegno sul tema: L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, Roma, Accademia Nazionale dei Lincei, 1968, pp. 219-226.

M. Bozzo, *La grande storia del computer: dall'abaco all'intelligenza artificiale*, Bari, Edizioni Dedalo, 1966, pp. 111-112.

V. Braitenberg, "Contributi tecnici e concettuali dell'elettronica allo studio dei cervelli viventi", in Aa. Vv., *Atti del convegno sul tema: L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, Roma, Accademia Nazionale dei Lincei, 1968, pp. 211-217.

S. Buffo, *Cibernetica e fantasmi*, 2011, consultabile presso il sito:
<http://www.paeseroma.it/wordpress/cibernetica-e-fantasmi-appunti-sulla-narrativa-come-processocombinatorio/>

G. Calogero, "Lettera a Vittorio Somenzi", in Aa. Vv., *L'uomo e la macchina. Atti del XXI Congresso Nazionale di Filosofia, Pisa 22-25 aprile 1967*, Torino, Edizioni di «Filosofia», 1967, vol. 2, pp. 94-96.

G. Calogero, "L'uomo, l'automa e lo schiavo", in *La cultura*, n. 1, 1966, pp. 1-11.

- I. Calvino, “Cibernetica e fantasmi. Appunti sulla narrativa come processo combinatorio”, 1967, in I. Calvino, *Saggi*, vol. I, Milano, Mondadori, 1995, pp. 205-226.
- I. Calvino, *Una pietra sopra. Discorsi di letteratura e società*, Torino, Einaudi, 1980.
- G. Capriz, “La matematica ed i calcolatori elettronici”, in Aa. Vv., *Atti del convegno sul tema: L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, Roma, Accademia Nazionale dei Lincei, 1968, pp. 55-67.
- P. Cariani, “On the Importance of Being Emergent”, in *Constructivist Foundations*, vol. 5, n. 2, 2009, pp. 86-91.
- S. Ceccato, “Intelligenza artificiale”, voce per l'*Enciclopedia della Scienza e della Tecnica*, Mondadori EST, 1965.
- S. Ceccato, “La grammatica insegnata alle macchine”, in *Civiltà delle macchine*, a. IV, n. 1-2, 1956, pp. 1-22.
- S. Ceccato, *La mente vista da un cibernetico*, ERI Edizioni, Torino, 1972.
- S. Ceccato, “La storia di un modello meccanico dell'uomo che traduce”, in Aa. Vv., *Almanacco Letterario Bompiani*, Milano, Bompiani, 1962, pp. 122-34.
- S. Ceccato, “Le prime accoglienze ad Adamo II”, in *Civiltà delle macchine*, a. IV, n. 3, 1956, p. 32.
- S. Ceccato, “Leonardo Sinisgalli, Civiltà delle macchine e Adamo II”, in Aa. Vv., *Atti del Simposio di studi su Leonardo Sinisgalli, Matera-Montemurro, 14-15-16 maggio 1982*, Matera, Liantonio, 1987, pp. 495-504. Consultabile anche presso il sito: <http://tysm.org/adamo-ii/>
- P. Churchland, V. Ramachandran, T. Sejnowski, “A Critique of Pure Vision”, in C. Koch, J. Davis, *Large-Scale Neuronal Theories of the Brain*, Cambridge, MIT press, 1994, pp. 23-60.
- F. Ciotti, G. Crupi, *Dall'Informatica umanistica alle culture digitali Atti del convegno di studi (Roma, 27-28 ottobre 2011) in memoria di Giuseppe Gigliozzi*, Roma, Quaderni Digilab - Casa Editrice Università La Sapienza, 2011, versione digitale disponibile presso il sito: http://digilab-epub.uniroma1.it/index.php/Quaderni_DigiLab/issue/view/2/showToc
- Civiltà delle Macchine, Leonardo Sinisgalli e Alan Turing:*
<http://matematica.unibocconi.it/articoli/civilt%C3%A0-delle-macchine-leonardo-sinisgalli-e-alanturing#1>

- H. M. Collins, "Stages in the Empirical Programme of Relativism", in *Social Studies of Science*, vol. 11, n.1, 1981, pp. 3-10.
- B. Copeland, *The Essential Turing*, Oxford, The Oxford University Press, 2004, p. 407.
- G. Corbellini, R. Cordeschi, "Natura, macchine, cervello e conoscenza. Attualità del pensiero di Vittorio Somenzi", in Aa. Vv., *Vittorio Somenzi 1918-2003: antologia e testimonianze*, Mantova, Fondazione Banca agricola mantovana, 2011, pp. 33-45.
- R. Cordeschi, "AI Turns Fifty: Revisiting its Origins", in *Applied Artificial Intelligence*, vol. 21, n. 3, 2007, pp. 259-279.
- R. Cordeschi, G. Tamburrini, "L'Intelligenza Artificiale: la storia e le idee", in E. Burattini, R. Cordeschi, *Intelligenza artificiale. Manuale per le discipline della comunicazione*, Roma, Carrocci, 2001, pp. 15-41.
- R. Cordeschi, "L'Intelligenza Artificiale", in L. Geymonat, *Storia del pensiero filosofico e scientifico*, vol. 8, t. III (a cura di E. Bellone e C. Mangione), Milano, Garzanti, 1996, pp. 145-200.
- R. Cordeschi, M. Frixione, "Premessa alla traduzione italiana del documento di Dartmouth", in *Sistemi Intelligenti*, vol. 18, n. 3, 2006, pp. 407-413.
- R. Cordeschi, *The Discovery of the Artificial: Behavior, Mind and Machines Before and Beyond Cybernetics*, Dordrecht, Kluwer Academic Publishers, 2002.
- R. Cordeschi, V. Somenzi, *La filosofia degli automi*, terza edizione, Torino, Boringhieri, 1986.
- J. D. Cowan, "Storia dei concetti e delle tecniche nella ricerca sulle reti neurali", in *Frontiere della Vita*, 1999: [http://www.treccani.it/enciclopedia/storia-dei-concetti-e-delle-tecniche-nella-ricerca-sulle-reti-neurali_\(Frontiere_della_Vita\)/](http://www.treccani.it/enciclopedia/storia-dei-concetti-e-delle-tecniche-nella-ricerca-sulle-reti-neurali_(Frontiere_della_Vita)/)
- E. T. Crawford, A. D. Biderman, *Social Scientists and International Affairs. A Case For A Sociology Of Social Science*, New York, Wiley, 1969.
- D. Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*, New York, Basic Books, 1993.
- A. Cuzzer, "La diffusione dell'informatica in Italia", in Aa. Vv., *La cultura informatica in Italia. Riflessioni e testimonianze sulle origini. 1950-1970*, Torino, Bollati Boringhieri - Fondazione Adriano Olivetti, 1993, pp. 3-36.
- R. De Benedetti, "Il calcolo proposizionale ovvero l'algebra delle idee", in Aa. Vv., *Almanacco Letterario Bompiani*, Milano, Bompiani, 1962, pp. 89-91.

- U. Eco, "La forma del disordine", in Aa. Vv., *Almanacco Letterario Bompiani*, Milano, Bompiani, 1962, pp. 175-187.
- P. Edwards, *Closed World: Computers and the Politics of Discourse in Cold War America*, Cambridge, MIT Press, 1996, pp. 339-372.
- P. Edwards, "The Machine in the Middle: Cybernetic Psychology and World War II", in R. Chrisley, S. Begeer, *Artificial Intelligence: Critical Concepts*, vol. I, Londra, Routledge, 2000, p. 338-371.
- J. Fodor, *The Language of Thought*, New York, Thomas Y. Crowell, 1975.
- g.j., "I filosofi, l'uomo e la macchina", in *Civiltà delle macchine*, a. XV, n. 3, 1967, pp. 2-6.
- S. W. Golomb, E. Berlekamp, T. M. Cover, R. G. Gallager, J. L. Massey, A. J. Viterbi, "Claude Elwood Shannon (1916–2001)", in *Notices of the American Mathematical Society*, vol. 49, n. 1, 2002, pp. 8-16.
- R. Granger, "Essential Circuits of Cognition: The Brain's Basic Operations", architecture, and representations, in J. Moor, G. Cybenko, *AI at 50: The future of Artificial Intelligence*, 2006, consultabile presso il sito:
<http://www.dartmouth.edu/~rhg/pubs/RHGai50.pdf>
- A. Guzzo, "L'uomo, la macchina, la tecnica", in Aa. Vv., *L'uomo e la macchina. Atti del XXI Congresso Nazionale di Filosofia, Pisa 22-25 aprile 1967*, Torino, Edizioni di «Filosofia», 1967, vol. 1, pp. 1-26.
- P. Hayes, J. McCarthy, "Some Philosophical Problems from the Standpoint of Artificial Intelligence", in B. Meltzer, D. Michie, *Machine Intelligence*, Edimburgo, Edinburgh University Press, 1969, pp. 463-502. Consultabile presso il sito:
<http://www-formal.stanford.edu/jmc/mchay69.pdf>
- P. Hayes, L. Morgenstern, "On John McCarthy's 80th Birthday, in Honor of His Contributions", in *AI Magazine*, vol. 28, n. 4, 2007, pp. 93-102.
- P. Hayes, "The Frame Problem and Related Problems", in B. L. Webber, N. J. Nilsson, *Readings in Artificial Intelligence*, Los Altos, Morgan Kaufmann, 1981, pp. 223-230.
- S. Heims, *I Cibernetici: Un gruppo e un'idea*, Roma, Editori Riuniti, 1994.
- A. Hodges, *Alan Turing: Una biografia*, Torino, Boringhieri, 2006.
- M. Houston, *AI@50: The Future of Learning and Search*, 2006a:
http://www.megmaker.com/2006/07/ai50_the_future_2.html

M. Houston, *AI@50: AI Past, Present, Future*, 2006b:
http://www.megmaker.com/2006/07/ai50_ai_past_pr.html

M. Houston, *AI@50: Future Interactions with Intelligent Machines*, 2006c:
http://www.megmaker.com/2006/07/ai50_future_int.html

M. Houston, *AI@50: Opening*, 2006d:
http://www.megmaker.com/2006/07/ai50_opening.html

M. Houston, *AI@50: The Future of AI*, 2006e:
http://www.megmaker.com/2006/07/ai50_the_future_3.html

M. Houston, *AI@50: The Future of Language and Cognition*, 2006f:
http://www.megmaker.com/2006/07/ai50_the_future_6.html

M. Houston, *AI@50: The Future of Network Models*, 2006g:
http://www.megmaker.com/2006/07/ai50_the_future_1.html

M. Houston, *AI@50: The Future of Vision*, 2006h:
http://www.megmaker.com/2006/07/ai50_the_future_4.html

P. Husbands, “An Interview with John Holland”, in P. Husbands, O. Holland, M. Wheeler, *The Mechanical Mind in History*, Londra, MIT Press, 2008, pp. 383-396.

P. Husbands, “An Interview with Oliver Selfridge”, in P. Husbands, O. Holland, M. Wheeler, *The Mechanical Mind in History*, Londra, MIT Press, 2008b, pp. 397-408.

P. Husbands, O. Holland, “The Ratio Club: A Hub of British Cybernetics”, in P. Husbands, O. Holland, M. Wheeler, *The Mechanical Mind in History*, Londra, MIT Press, 2008, pp. 90-148.

W. J. Hutchins, *Machine translation: past, present, future*, New York, Halsted Press, 1986.

IBM Archives, alcuni dei fatti più significativi del 1956: https://www-03.ibm.com/ibm/history/history/year_1956.html

IBM Archives, approfondimento sui creatori dell'IBM 701: https://www-03.ibm.com/ibm/history/exhibits/701/701_team.html

R. R. Kline, “Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence”, in *IEEE Annals of the History of Computing*, vol. 33, n. 4, 2011, pp. 5-16.

R. R. Kline, *The Cybernetics Moment, Or Why We Call Our Age the Information Age*, Baltimora, Johns Hopkins University Press, 2015.

- S. Knapp, *Artificial Intelligence: Past, Present, and Future*, 2006a: www.dartmouth.edu/~vox/0607/0724/ai50.html
- S. Knapp, *Dartmouth receives grant from DARPA to support AI@50 conference*, 2006b: <http://www.dartmouth.edu/~news/releases/2006/07/06.html>
- S. Knapp, *July Conference at Dartmouth Commemorates Golden Anniversary of "Artificial Intelligence"*, 2006c: www.dartmouth.edu/~news/releases/2006/03/13a.html
- D. M. Li, P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*, New York, Springer Science & Business Media, 2013, p. 79.
- F. Lucentini, "Automatopoietica", in Aa. Vv., *Almanacco Letterario Bompiani*, Milano, Bompiani, 1962, pp. 152-158.
- A. K. Mackworth, "AI@50: We Are Golden!", in *AI Magazine*, vol. 27, n. 4, 2006, pp. 17-18.
- V. Marchis, "Quale futuro per l'industria italiana?", in *Il Contributo italiano alla storia del Pensiero*, 2013: [http://www.treccani.it/enciclopedia/quale-futuro-per-l-industria-italiana_\(Il_Contributo_italiano_alla_storia_del_Pensiero:_Tecnica\)/](http://www.treccani.it/enciclopedia/quale-futuro-per-l-industria-italiana_(Il_Contributo_italiano_alla_storia_del_Pensiero:_Tecnica)/)
- E. Maretti, "Adamo II", in *Civiltà delle macchine*, a. IV, n. 3, 1956, pp. 25-32;
- J. McCarthy, "Programs with Common Sense", in *Mechanisation of Thought Processes: Proceedings of a Symposium Held at the National Physical Laboratory on 24th, 25th, 26th and 27th November 1958*, vol. I, Londra, Her Majesty's Stationery Office, 1959, pp. 75-84. Consultabile presso il sito: <http://www-formal.stanford.edu/jmc/mcc59.pdf>
- J. McCarthy, "Circumscription - A Form of Non-Monotonic Reasoning", in *Artificial Intelligence*, vol. 13, n. 1-2, 1980, pp. 27-39. Consultabile presso il sito: <http://i.stanford.edu/pub/ctr/reports/cs/tr/80/788/CS-TR-80-788.pdf>
- J. McCarthy, lettera a Robert S. Morison, 8 dicembre 1955, consultabile presso il sito: <http://rockefeller100.org/files/original/d71f117b5e3d69c2cf160916e5821313.pdf>
- J. McCarthy, lettera a Robert S. Morison, 25 maggio 1956, consultabile presso il sito: <http://rockefeller100.org/files/original/48e0e29c4c63b94a6e87fbbd9ef150ba.pdf>
- J. McCarthy, "Review of The Question of Artificial Intelligence", in *Annals of the History of Computing*, 1989, consultabile presso il sito: <http://www-formal.stanford.edu/jmc/reviews/bloomfield.pdf>

- J. McCarthy, “The Future of AI: A Manifesto”, in *AI Magazine*, vol. 26, n. 4, 2006, pp. 39.
- J. McCarthy, “The Inversion of Functions Defined by Turing Machines”, in C. Shannon, J. McCarthy, *Automata Studies*, Princeton University Press, 1956, pp. 177–181. Consultabile presso il sito:
<http://jmc.stanford.edu/articles/inversion/inversion.pdf>
- J. McCarthy, C. E. Shannon, “Automata Studies: Preface”, 1956, ristampato in N. Sloane, A.D. Wynerm, *Claude Elwood Shannon: Collected Papers*, Piscataway, IEEE Press, 1993, pp. 727-732.
- J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon, *Proposal for the Dartmouth summer research project on artificial intelligence*, 1955, consultabile presso il sito:
<http://rockefeller100.org/files/original/ab461efeeb9ca28fd3de943abdd30b00.pdf>
- P. McCorduck, *Storia dell'intelligenza artificiale: gli uomini, le idee, le prospettive*, Padova, Franco Muzio, 1987.
- W. McCulloch, A. Oettinger, N. Rochester, O. Schmitt, “The design of machines to simulate the behavior of the human brain”, in *IRE transactions on Electronic Computers*, EC-5, 1956, pp. 240-255.
- W. McCulloch, W. Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity”, in *The bulletin of mathematical biophysics*, vol. 5, n. 4, 1943, pp 115-137.
- D. Mckay, “Mindlike Behaviour in Artefacts”, in *British Journal for the Philosophy of Science*, vol. 2, n. 6, 1951, pp. 105-121.
- D. A. Mindell, *Between Human and Machine: Feedback, Control, and Computing Before Cybernetics*, Baltimora, The Johns Hopkins University Press, 2002, p. 107.
- M. L. Minsky, “Memoir on Inventing the Confocal Scanning Microscope”, in *Scanning*, n. 10, 1988, pp. 128-38.
- M. L. Minsky, S. Papert, *Perceptrons. An Introduction to Computational Geometry*, Cambridge. M.I.T. Press, 1969.
- M. L. Minsky, *Semantic Information Processing*, Cambridge, MIT Press, 1968.
- M. L. Minsky, “Some Methods Of Artificial Intelligence And Heuristic Programming”, in *Proceeding of the Teddington Symposium on Mechanisation of Thought Processes*, vol. 1, 1959, pp. 3-27.
- M. L. Minsky, “Steps Toward Artificial Intelligence”, in *Proceedings of the IRE*, vol. 49, n. 1, 1961, pp. 8-30.

G. Montalenti, “III seduta. 17 ottobre 1967”, in Aa. Vv., *Atti del convegno sul tema: L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, Roma, Accademia Nazionale dei Lincei, 1968, pp. 163-164.

J. Moor, “The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years”, in *AI Magazine*, vol. 27, n. 4, 2006, pp. 87-91.

R. Mordenti, *Informatica e critica dei testi*, Roma, Bulzoni, 2001.

M. Morelli, *Dalle calcolatrici ai computer degli anni Cinquanta*, Milano, Franco Angeli, 2001.

R. S. Morison, impressioni su John McCarthy and Claude Shannon, 17 giugno 1955, consultabile presso il sito:
<http://rockefeller100.org/files/original/727608454ce30c7a24c15ef0ffd3bdc6.pdf>

R. S. Morison, lettera a John McCarthy, 30 novembre 1955, consultabile presso il sito:
<http://rockefeller100.org/files/original/8046166e880b0e25a5d1510ca94d6c53.pdf>

R. S. Morrison, lettera a Warren Weaver, 6 April 1955, consultabile presso il sito:
<http://rockefeller100.org/files/original/aa2a5de6eba33472a9310cd67f4d3b5d.pdf>

A. Newell, H. A. Simon, “Computer Simulation of Human Thinking”, in *Science*, vol. 134, n. 3495, 1961, pp. 2011-2017.

A. Newell, H. A. Simon, “Computers in Psychology”, in R. D. Luce, R. R. Bush, E. Galanter, *Handbook of Mathematical Psychology*, vol. 1, New York, John Wiley & Sons, 1963, pp. 361-428.

A. Newell, H. A. Simon, “Models: Their Use and Limitations”, in L. White, *The State of the Social Sciences*, Chicago, University of Chicago Press, 1955, pp. 68–83. Consultabile presso il sito:
<http://digitalcollections.library.cmu.edu/portal/main.jsp?flag=browse&smd=1&awdid=1>

A. Newell, J. Shaw, H. A. Simon, “Report On A General Problem Solving Program”, in *Proceedings of the International Conference on Information Processing*, Parigi, UNESCO, 1958, pp. 256-264. Consultabile per il sito:
http://bitsavers.informatik.uni-stuttgart.de/pdf/rand/ipl/P-1584_Report_On_A_General_Problem-Solving_Program_Feb59.pdf

A. Newell, “The First AAAI President’s Message”, in *AI Magazine*, vol. 26, n. 4, 2005, pp. 24-29.

N. J. Nilsson, *The Quest For Artificial Intelligence. A History Of Ideas And Achievements*, Cambridge University Press, 2009, versione digitale disponibile presso il sito: <http://ai.stanford.edu/~nilsson/QAI/qai.pdf>

A. Norberg, *An Interview with Allen Newell*, Pittsburgh, Charles Babbage Institute, 1991, consultabile presso il sito:
<https://conservancy.umn.edu/bitstream/handle/11299/107544/oh227an.pdf?sequence=1>

A. Norberg, *An Interview with Marvin L. Minsky*, Cambridge, Charles Babbage Institute, 1989, consultabile presso il sito:
<https://conservancy.umn.edu/bitstream/handle/11299/107503/oh179mlm.pdf?sequence=1>

A. Oettinger, “Programming a Digital Computer to Learn”, in *Philosophical Magazine*, vol. 43, 1952, pp. 1243-1263.

M. Olazaran, “A Sociological Study of the Official History of the Perceptrons Controversy”, in *Social Studies of Science*, vol. 26, n. 3, 1996, pp. 611–659

“Omeòstato”, in *Enciclopedia on line Treccani*:
<http://www.treccani.it/vocabolario/omeostato/>

M. Pacifico, “I nuovi Gutenberg”, in Aa. Vv., *Almanacco Letterario Bompiani*, Milano, Bompiani, 1962, pp.100-103.

D. Parisi, “Reti neurali e vita artificiale”, in *Enciclopedia della Scienza e della Tecnica*, 2007:
[http://www.treccani.it/enciclopedia/reti-neurali-e-vita-artificiale_\(Enciclopedia-della-Scienza-e-della-Tecnica\)/](http://www.treccani.it/enciclopedia/reti-neurali-e-vita-artificiale_(Enciclopedia-della-Scienza-e-della-Tecnica)/)

W. Pitts, “Comments on Session on Learning Machines”, in *Proceedings of the March*, n. 1-3, 1955, pp. 108-111.

C. Pogliano, “1967: il congresso della Società Filosofica Italiana. Cronaca delle quattro giornate pisane”, in *I quaderni di Athenet*, n. 20, 2007, pp. 4-8.

C. Pogliano, “Alla periferia del nascente Impero: il caso Italia (1945-1968)”, in *Discipline Filosofiche*, a. XVII, n. 1, 2007, pp. 85-120.

C. Pogliano, “Le nuove macchine: inquietudine e seduzione”, in F. Cassata, C. Pogliano, *Storia d'Italia, Annali 26, Scienze e cultura nell'Italia Unita*, Torino, Einaudi, 2011, pp. 349-381.

C. Pogliano, “Sciences at War and the Cybernetic Dream”, in *Nuncius. Annali di storia della scienza*, vol. 19, n. 1, 2004, pp.171-204.

“Ricerca operativa”, in *Enciclopedia on line Treccani*:
<http://www.treccani.it/enciclopedia/ricerca-operativa/>

E. Rich, *Intelligenza artificiale*, Milano, McGraw-Hill, 1986.

- P. Ridolfi, *Cinquant'anni con i BIT*, Roma, 2007, versione digitale disponibile presso il sito: http://archivio.cnipa.gov.it/site/_files/libro_ridolfi_2007.pdf
- S. Rose, *The Making of Memory: From Molecules to Mind*, New York, Anchor Books/Doubleday, 1993.
- F. Rosenblatt, "Perceptron Simulation Experiments", in *Proceedings of the IRE*, vol. 48, n. 3, 1960, pp. 301-309.
- A. Rosenblueth, N. Wiener, J. Bigelow, "Behavior, Purpose and Teleology", in *Philosophy of Science*, vol. 10, n. 1, 1943, pp. 18-24.
- S. J. Russell, P. Norvig, *Intelligenza Artificiale: Un Approccio Moderno*, vol. 2, Milano, Pearson Education, 2005, p. 547.
- S. Salvia, "L'uomo e la macchina. Passato e presente", in *Rivista di Storia della Filosofia*, a. 2009, fasc. 3, pp. 505-11.
- J. E. Sammet, "Programming Languages: History and Future", in *Communications of the ACM*, vol. 15, n. 7, 1972, pp. 601-610.
- J. R. Searle, "Minds, Brains, and Programs", in *Behavioral and Brain Sciences*, vol. 3, n. 3, 1980, pp. 417-457.
- B. Segre, "Saluto", in Aa. Vv., *Atti del convegno sul tema: L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, Roma, Accademia Nazionale dei Lincei, 1968, pp. 9-12.
- O. G. Selfridge, "Pandemonium: A Paradigm for Learning", in D. V. Blake, A. M. Uttley, *Mechanisation of Thought Processes*, vol. I, Londra, Her Majesty's Stationery Office, 1959, pp. 511-529.
- O. G. Selfridge, "Pattern Recognition and Modern Computers", in *Proceedings of the 1955 Western Joint Computer Conference*, n. 1-3, 1955, pp. 91-93.
- C. E. Shannon, "A Universal Turing Machine with Two Internal States", 1954, ristampato in N. Sloane, A.D. Wyner, *Claude Elwood Shannon: Collected Papers*, Piscataway, IEEE Press, 1993, pp. 733-741.

- C. E. Shannon, "Computers and Automata", 1953, ristampato in N. Sloane, A.D. Wyner, *Claude Elwood Shannon: Collected Papers*, Piscataway, IEEE Press, 1993, pp. 703-711.
- C. E. Shannon, "Presentation of a Maze Solving Machine", 1951, ristampato in N. Sloane, A.D. Wyner, *Claude Elwood Shannon: Collected Papers*, Piscataway, IEEE Press, 1993, pp. 681-688.
- C. E. Shannon, "Programming a Computer for Playing Chess", 1950, ristampato in N. Sloane, A.D. Wyner, *Claude Elwood Shannon: Collected Papers*, Piscataway, IEEE Press, 1993, pp. 637-657.
- H. A. Simon, "Modeling Human Mental Processes", in *Proceeding of the Western Joint Computer Conference*, vol. 19, 1961, pp. 111-120.
- A. H. Simon, *Models of My Life*, New York, Basic Books, 1991.
- A. H. Simon, "Models: Their uses and limitations", in L. D. White, *The State of Social Science*, Chicago, Chicago University Press, 1956, pp. 66-83.
- A. H. Simon, "The future of information systems", in *Annals of Operation Research*, vol. 71, n. 0, 1997, pp. 3-14.
- J. Simonsen, *Herbert A. Simon: Administrative Behavior. How Organizations Can Be Understood in Terms of Decision Processes*, Roskilde University, 1994, consultabile presso il sito: <http://jespersimonsen.dk/Downloads/Simon-introduction.pdf>
- L. Sinisgalli, "Biblioteca", in *Civiltà delle Macchine*, a. I, n. 2, marzo 1953a, p. 80.
- L. Sinisgalli, "Semaforo", *Civiltà delle macchine*, a. I, n. 3, luglio 1953b, p. 79.
- N. Sloane, A. D. Wyner, *Claude Elwood Shannon: Collected Papers*, Piscataway, IEEE Press, 1993.
- V. Somenzi, "Bigliografia sul Congresso", in Aa. Vv., *L'uomo e la macchina. Atti del XXI Congresso Nazionale di Filosofia, Pisa 22-25aprile 1967*, Torino, Edizioni di «Filosofia», 1967, vol. 3, pp. 304-310.
- V. Somenzi, "Uomini e macchine", in Aa. Vv., *L'uomo e la macchina. Atti del XXI Congresso Nazionale di Filosofia, Pisa 22-25aprile 1967*, Torino, Edizioni di «Filosofia», 1967, vol. I, pp. 51-65.

V. Somenzi, “Aspetti filosofici del problema della «intelligenza artificiale»”, in Aa. Vv., *Atti del convegno sul tema: L'automazione elettronica e le sue implicazioni scientifiche, tecniche e sociali*, Roma, Accademia Nazionale dei Lincei, 1968, pp. 227-232.

V. Somenzi, “Cibernetica, informatica e filosofia della scienza”, in Aa. Vv., *La cultura informatica in Italia. Riflessioni e testimonianze sulle origini. 1950-1970*, Torino, Boringhieri - Fondazione Adriano Olivetti, 1993, pp. 161-190.

V. Somenzi, *La filosofia degli automi*, Torino, Boringhieri, 1965.

G. Spezzano, D. Talia, *Calcolo parallelo, automi cellulari e modelli per sistemi complessi*, Milano, Franco Angeli, 1999, p. 144.

A. M. Turing, “Computing Machinery and Intelligence”, in *Mind*, vol. 49, n. 236, 1950: 433-460.

A. M. Turing, *Intelligent Machinery*, 1948, consultabile presso il sito:
http://www.alanturing.net/intelligent_machinery/

A. M. Turing, “On Computable Numbers, with an Application to the Entscheidungsproblem”, in *Proceedings of the London Mathematical Society*, ser. 2, vol. 42, 1936, pp. 230-265.

A. M. Turing, “The Chemical Basis of Morphogenesis”, in *Philosophical Transactions of the Royal Society of London*, vol. 237, n. 641, 1952, pp. 37-72.

R. Urbani, “L'uomo e la macchina. Un congresso di trent'anni fa”, in *Nuncius. Annali di storia della scienza*, vol. 12, n. 2, 1967, pp. 497-524.

J. von Neumann, “Probabilistic Logics and the Synthesis of Reliable Organisms From Unreliable Components”, in C. Shannon, J. McCarthy, *Automata Studies*, Princeton University Press, 1956, pp. 43-98.

W. McCulloch, “What is a Number, that a Man May Know It, and a Man, that He May Know a Number? Alfred Korzybski Memorial Lecture”, in *General Semantics Bulletin*, n. 26-27, 1960, pp. 7-18.

W. Weaver, impressioni su McCarthy tratti dalla sua agenda personale, 4 aprile 1955, consultabile presso il sito:
<http://rockefeller100.org/files/original/f0ce90fc3583b6d211eefaa075674227.pdf>

W. Weaver, memorandum su John McCarthy and Claude Shannon, 14 giugno 1955, consultabile presso il sito:
<http://rockefeller100.org/files/original/0469561e6ff96a1541cd41377687e098.pdf>

W. G. Wiederhold, J. McCarthy, E. Feigenbaum. *Memorial Resolution: Arthur L. Samuel (1901 - 1990)*, 1990, consultabile presso il sito:
<http://historicalsociety.stanford.edu/pdfmem/SamuelA.pdf>

N. Wiener, *Cybernetics or control and communication in the animal and the machine*, New York, John Wiley & Sons, 1949.

N. Wiener, *I Am a Mathematician*, New York, Doubleday & Company, 1956.

N. Wiener, "Some Moral and Technical Consequences of Automation", in *Science (New Series)*, vol. 131, n. 3410, 1960, pp. 1355-1358.

M. Wilkes, "Can Machines Think?", in *Proceeding of the IRE*, vol. 41, n. 10, 1953, pp. 1230-1234.

L. Zanetti, "Bompiani ordina poesia a macchina", in *L'espresso*, 10 dicembre 1961, una scansione parziale dell'articolo è consultabile presso il sito:
http://www.edixxon.com/computerstory/immagini/pgn_1110_115.html